

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2002, 10:44:42 ; Search time 10.35 seconds

(without alignments)
59.856 Million cell updates/sec

Title: US-09-674-254-3

Perfect score: 87

Sequence:

1 CASDGLVPRRLQHRP 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	77.0	366	1	LFTB_HUMAN
2	67	77.0	366	1	TGF4_HUMAN
3	66	75.9	368	1	TGF4_MOUSE
4	52	59.8	368	1	LFTB_MOUSE
5	45	51.7	293	1	TRPB_ECOLI
6	44	50.6	293	1	TRPB_SALTY
7	43	49.4	274	1	TRPB_HAEN
8	42	48.3	335	1	HYPB_RHOCA
9	40.5	46.6	351	1	RPBB_XANCP
10	40	46.0	499	1	OPSL_PATYE
11	40	46.0	695	1	SPAL_SYNY3
12	40	46.0	847	1	NIRB_ECOLI
13	40	46.0	1204	1	NOS3_BOVIN
14	39	44.8	248	1	ATPI_PORPU
15	39	44.8	345	1	OGG1_HUMAN
16	39	44.8	448	1	FIBP_ADEP3
17	39	44.8	650	1	Y411_HUMAN
18	38.5	44.3	919	1	NOS3_RAT
19	38.5	44.3	1201	1	NOS3_MOUSE
20	38.5	44.3	1204	1	NOS3_PIG
21	38	43.7	141	1	YM2_DHTV1
22	38	43.7	251	1	ATP6_ANASP
23	38	43.7	351	1	WNT4_HUMAN
24	38	43.7	351	1	WNT4_MOUSE
25	38	43.7	351	1	WNT4_RAT
26	38	43.7	383	1	FLHB_YEREN
27	38	43.7	439	1	PTV2_ADEG1
28	38	43.7	506	1	VE12_HPV16
29	38	43.7	825	1	5E5_RAT
30	38	43.7	4753	1	LRP_CAEEL
31	37.5	43.1	226	1	MINC_BACSU
32	37.5	43.1	1202	1	NOS3_HUMAN
33	37	42.5	75	1	COXH_RAT

34	37	42.5	137	1	RL16_STRPN
35	37	42.5	140	1	YDAT_ECOLI
36	37	42.5	184	1	YTFJ_ECOLI
37	37	42.5	219	1	RL10_BOMMA
38	37	42.5	233	1	ATPI_OCHNE
39	37	42.5	288	1	HEMK_MYCLE
40	37	42.5	346	1	TBCC_HUMAN
41	37	42.5	395	1	GP44_HUMAN
42	37	42.5	659	1	SPA3_SYNY3
43	37	42.5	1188	1	YEC5_YEAST
44	37	42.5	1331	1	CYAB_LEIDO
45	36	41.4	58	1	YBAS_PHAVU

ALIGNMENTS

RESULT 1
LFTB_HUMAN
ID LFTB_HUMAN STANDARD; PRT; 366 AA.
AC 075610;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Left-right determination factor B precursor (Lefty-B protein).
GN LFTB OR LEFTYB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Teratocarcinoma;
RX MEDLINE=99162193; PubMed=10053005;
RA Kosaki K., Bassi M.T., Kosaki R., Lewin M., Belmont J., Schauer G.,
RA Casey B.;
RT "Characterization and mutation analysis of human LEFTY A and LEFTY B,
RT homologues of murine genes implicated in left-right axis
RT development";
RL Am. J. Hum. Genet. 64:712-721(1999).
CC -!- FUNCTION: REQUIRED FOR LEFT-RIGHT ASYMMETRY DETERMINATION OF
CC ORGAN SYSTEMS IN MAMMALS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: THE PROCESSING OF THE PROTEIN MAY ALSO OCCUR AT THE SECOND R-
CC X-X-R SITE LOCATED AT AA 132-135. PROCESSING APPEARS TO BE
CC REGULATED IN A CELL-TYPE SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF081507; AAC33967.1; JOINED.
EMBL; AF081504; AAC33967.1; JOINED.
EMBL; AF081505; AAC33967.1; JOINED.
EMBL; AF081506; AAC33967.1; JOINED.
EMBL; AF081512; AAC48144.1; JOINED.
HSP; F10600; ITG3.
MIM; 603037; -.
InterPro; IPR001839; TGF-beta.
InterPro; IPR001111; TGF-beta.
Pfam; PF00019; TGF-beta; 1.
Pfam; PF00688; TGF-beta; 1.
ProDom; PD000357; TGF-beta; 1.
SMART; SM00204; TGF-beta; 1.
PROSITE; PS00250; TGF-beta; 1.
Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal;
Multigene family.
POTENTIAL.

1 21

FT PROPEP 22 76 OR 135 (POTENTIAL).
 FT CHAIN 77 366 LEFT-RIGHT DETERMINATION FACTOR B.
 FT DISULFID 251 264 BY SIMILARITY.
 FT DISULFID 263 316 BY SIMILARITY.
 FT DISULFID 293 351 BY SIMILARITY.
 FT DISULFID 297 353 BY SIMILARITY.
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 366 AA: 40880 MW: BCF900C71ED9AA2A CRC64;

Query Match 77.0%; Score 67; DB 1; Length 366;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQ 13
 b 353 CASDGLVPRRLQ 365
 |||||

RESULT 2
 TGF4_HUMAN STANDARD; PRT: 366 AA.
 ID TGF4_HUMAN STANDARD; PRT: 366 AA.
 AC 000292; 075611;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Transforming growth factor beta 4 precursor (TGF-beta 4) (Endometrial
 bleeding-associated factor) (left-right determination factor A)
 DE (Lefty-A protein).
 GN EBAF OR TGFb4 OR LEFTA OR LEFTYA.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97298127; PubMed=9153275;
 RA Kothapalli R., Buyuksal I., Wu S.-Q., Chegini N., Tabibzadeh S.;
 RT "Detection of ebafr, a novel human gene of the transforming growth
 factor beta superfamily association of gene expression with
 endometrial bleeding.";
 RL J. Clin. Invest. 99:2342-2350(1997).
 RN [2]
 :P REVISIONS.
 RX MEDLINE=99162193; PubMed=10053005;
 RA Kothapalli R.;
 RL Unpublished results, cited by:
 RL Kosaki K., Bassi M.T., Kosaki R., Lewin M., Belmont J., Schauer G.,
 RL Casey B.;
 RL Am. J. Hum. Genet. 64:712-721(1999).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT L-R AXIS MALFORMATIONS ASN-342.
 RC TISSUE=Placenta;
 RX MEDLINE=99162193; PubMed=10053005;
 RA Kosaki K., Bassi M.T., Kosaki R., Lewin M., Belmont J., Schauer G.,
 RA Casey B.;
 RT "Characterization and mutation analysis of human LEFTY A and LEFTY B,
 RT homologues of murine genes implicated in left-right axis
 RT development.";
 RL Am. J. Hum. Genet. 64:712-721(1999).
 CC -1- FUNCTION: REQUIRED FOR LEFT-RIGHT ASYMMETRY DETERMINATION OF
 CC ORGAN SYSTEMS IN MAMMALS. MAY PLAY A ROLE IN ENDOMETRIAL BLEEDING.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: MESENCHYMAL CELLS OF THE ENDOMETRIAL STROMA.
 CC -1- DEVELOPMENTAL STAGE: TRANSIENTLY EXPRESSED BEFORE AND DURING
 CC MENSTRUAL BLEEDING.
 CC -1- PTM: THE PROCESSING OF THE PROTEIN MAY ALSO OCCUR AT THE SECOND R-
 CC X-X-R SITE LOCATED AT AA 132-135. PROCESSING APPEARS TO BE
 CC REGULATED IN A CELL-TYPE SPECIFIC MANNER.
 CC -1- DISEASE: DEFECTS IN EBAF RESULT IN LEFT-RIGHT AXIS MALFORMATIONS
 CC INCLUDING LEFT PULMONARY ISOMERISM, CARDIAC ANOMALIES
 CC CHARACTERIZED BY COMPLETE ATRIOVENTRICULAR CANAL DEFECT AND

CC HYPOPLASTIC LEFT VENTRICLE, AND INTERRUPTED INFERIOR VENA CAVA.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U81523; AAB53269.1; ALT_SEQ.
 DR EMBL; AF081511; AAC32600.1;
 DR EMBL; AF081508; AAC32600.1; JOINED.
 DR EMBL; AF081509; AAC32600.1; JOINED.
 DR EMBL; AF081510; AAC32600.1; JOINED.
 DR EMBL; AF081513; AAD48145.1;
 DR HSSP; P10600; ITGJ.
 DR MIM; 601877;
 DR InterPro; IPR001839; TGF-beta.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 2.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR ProDom; PD000357; TGF-beta; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal;
 KW Multigene family; Disease mutation.
 FT SIGNAL 1 21
 FT PROPEP 22 76
 FT CHAIN 77 366
 FT DISULFID 251 264
 FT DISULFID 263 316
 FT DISULFID 293 351
 FT DISULFID 297 353
 FT CARBOHYD 158 158
 FT VARIANT 342 342
 FT SEQUENCE 366 AA: 40920 MW: 63A416CAE30F7A39 CRC64;
 /FTID-VAR_010385

Query Match 77.0%; Score 67; DB 1; Length 366;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQ 13
 Db 353 CASDGLVPRRLQ 365
 |||||

RESULT 3
 TGF4_MOUSE STANDARD; PRT: 368 AA.
 ID TGF4_MOUSE STANDARD; PRT: 368 AA.
 AC 064280;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transforming growth factor beta 4 precursor (TGF-beta 4) (Lefty
 DE protein) (Lefty-1 protein) (STRA3 protein).
 GN EBAF OR TGFb4 OR STRA3 OR LEFTY OR LEFTY1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96202359; PubMed=8610011;
 RA Meno C., Saijoh Y., Fujii H., Ikeda M., Yokoyama T., Yokoyama M.,
 RA Toyoda Y., Hamada H.;
 RT "Left-right asymmetric expression of the TGF beta-family member lefty
 RT in mouse embryos.";
 RL Nature 381:151-155(1996).
 RN [2]

RP SEQUENCE FROM N.A.
RA Bouillet P.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RX SEQUENCE FROM N.A.
RX MEDLINE-98156497; PubMed-9496783;
RA Ould-Abdelghani M., Chazaud C., Bouillet P., Mattei M.-G., Dolle P.,
RA Chambon P.;
RT "Stra3/lefty, a retinoic acid-inducible novel member of the
RT transforming growth factor-beta superfamily.";
RL Int. J. Dev. Biol. 42:23-32(1998).
RN [4]
RP FUNCTION:
RX MEDLINE-98372436; PubMed-9708731;
RX Meno C., Shimono A., Saijoh Y., Yashiro K., Mochida K., Ohishi S.,
RA Noji S., Kondoh H., Hamada H.;
RA "Lefty-1 is required for left-right determination as a regulator of
RT Lefty-2 and nodal";
RL Cell 94:287-297(1998).
CC -!- FUNCTION: REQUIRED FOR LEFT-RIGHT AXIS DETERMINATION AS A
CC REGULATOR OF LEFT2 AND NODAL.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: BY E8.0, EXPRESSED EXCLUSIVELY ON THE LEFT
CC SIDE OF DEVELOPING EMBRYOS WITH EXPRESSION PREDOMINANTLY IN THE
CC PROSPECTIVE FLOOR PLATE (PFP). WEAK EXPRESSION IN THE LATERAL-
CC PLATE MESODERM (LPM).
CC -!- PTM: THE PROCESSING OF THE PROTEIN MAY ALSO OCCUR AT THE SECOND R-
CC X-X-R SITE LOCATED AT AA 132-135. PROCESSING APPEARS TO BE
CC REGULATED IN A CELL-TYPE SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D83921; BAA12121.1; -;
DR EMBL; 273151; CAA97497.1; -;
DR EMBL; AJ000082; CAA03909.1; -;
DR EMBL; AJ000083; CAA03910.1; -;
DR HSSP; P10600; ITGK.
DR MGD; MGI:107405; Ebaf.
DR InterPro; IPR001839; TGF-beta.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR ProDom; PD000357; TGF-beta; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF-BETA_1; 1.
DR Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal.
KW SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 76 OR 135 (POTENTIAL).
FT CHAIN 77 368 TRANSFORMING GROWTH FACTOR BETA 4.
FT DISULFID 253 366 BY SIMILARITY.
FT DISULFID 265 318 BY SIMILARITY.
FT DISULFID 295 353 BY SIMILARITY.
FT DISULFID 299 355 BY SIMILARITY.
FT CARBOHYD 158 158 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 368 AA; 41497 MW; 821DAE663C546B5F CRC64;

Query Match 75.9%; Score 66; DB 1; Length 368;
Best Local Similarity 92.3%; Pred. No. 0.00034;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQ 13
DB 355 CASDGLVPRRLQ 367

RESULT 4
LFTB_MOUSE
ID LFTB_MOUSE STANDARD; PRT; 368 AA.
AC P57785;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Left-right determination factor B precursor (Lefty-2 protein).
GN LEFTB OR LEFTY2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98006264; PubMed-9348041;
RA Meno C., Ito Y., Saijoh Y., Matsuda Y., Tashiro K., Kuhara S.,
RA Hamada H.;
RT "Two closely-related left-right asymmetrically expressed genes, lefty-
RT 1 and lefty-2; their distinct expression domains, chromosomal linkage
RT and direct neutralizing activity in Xenopus embryos.";
RL Genes Cells 2:513-524(1997).
CC -!- FUNCTION: MORPHOGEN FOR LEFT-RIGHT ASYMMETRY DETERMINATION OF
CC ORGAN SYSTEMS IN MAMMALS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: AT THE PRIMITIVE STREAK STAGE (E7.0),
CC EXPRESSED IN THE EMERGING MESODERM. BY E8.0, EXPRESSED EXCLUSIVELY
CC ON THE LEFT SIDE OF DEVELOPING EMBRYOS WITH EXPRESSION
CC PREDOMINANTLY IN THE LATERAL-PLATE MESODERM (LPM). WEAK EXPRESSION
CC IN THE PROSPECTIVE FLOOR PLATE (PFP).
CC -!- PTM: THE PROCESSING OF THE PROTEIN MAY ALSO OCCUR AT THE SECOND R-
CC X-X-R SITE LOCATED AT AA 132-135. PROCESSING APPEARS TO BE
CC REGULATED IN A CELL-TYPE SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR HSSP; P10600; ITGK.
DR MGD; MGI:1277956; Leftb.
DR InterPro; IPR001839; TGF-beta.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 2.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR ProDom; PD000357; TGF-beta; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF-BETA_1; 1.
DR Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 77 OR 135 (POTENTIAL).
FT CHAIN 78 368 LEFT-RIGHT DETERMINATION FACTOR B.
FT DISULFID 253 266 BY SIMILARITY.
FT DISULFID 265 318 BY SIMILARITY.
FT DISULFID 295 353 BY SIMILARITY.
FT DISULFID 299 355 BY SIMILARITY.
FT CARBOHYD 158 158 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 368 AA; 41141 MW; 56B5ED095167A6C3 CRC64;

Query Match 59.8%; Score 52; DB 1; Length 368;
Best Local Similarity 90.0%; Pred. No. 0.1;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPR 10
DB 355 CASDGLVPR 364

RESULT 5
TRPH_ECOLI
ID TRPH_ECOLI STANDARD; PRT; 293 AA.
AC P77766;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein trpH.

```

GN TRPH OR B1266.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12."
RA Science 277:1453-1474(1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map."
RL DNA Res. 3:363-377(1996).
CC
CC -!- SIMILARITY: BELONGS TO THE TRPH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000224; AAC74348.1;
CC EMBL; D90764; BAA14800.1;
CC EMBL; D90765; BAA14815.1;
CC EcoGene; EG14257; trph.
CC InterPro; IPR004013; PHP_C.
CC InterPro; IPR003141; PHP_N.
CC Pfam; PF02811; PHP_C; 1.
CC Pfam; PF02231; PHP_N; 1.
CC SMART; SM00481; POLIIC; 1.
CC Complete proteome.
CC SEQUENCE 293 AA; 32580 MW; 0A045925C472A15F CRC64;

Query Match 51.7%; Score 45; DB 1; Length 293;
Best Local Similarity 64.3%; Pred. No. 1.4; Mismatches 5; Indels 0; Gaps 0;
Matches 9; Conservative 0;

Qy 2 ASDGALVPRRLQHR 15
Db 18 ASDGCLTPEALVHR 31
|||||
18 ASDGCLTPEALVHR 31

RESULT 6
TRPH_SALTY
ID TRPH_SALTY STANDARD; PRT; 293 AA.
AC O54453;
DT 30-MAY-2000 (Rel. 39, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Protein trph.
GN TRPH OR STM1721.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;

```

```

RN
RP SEQUENCE FROM N.A.
RA Yang W., Bai Q., Skrypka I., Zhao G., Somerville R.L.;
RT "The 5'-upstream regions of the trp operons of Escherichia coli and
RT Salmonella typhimurium contain divergently oriented, TrpR-controlled
RT transcriptional units."
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT Lt2."
RL Nature 413:852-856(2001).
CC
CC -!- SIMILARITY: BELONGS TO THE TRPH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U92714; AAB93886.1; ALT_INIT.
CC EMBL; AE008776; AAL20639.1;
CC StyGene; SG10727; trph.
CC InterPro; IPR004013; PHP_C.
CC InterPro; IPR003141; PHP_N.
CC Pfam; PF02811; PHP_C; 1.
CC Pfam; PF02231; PHP_N; 1.
CC SMART; SM00481; POLIIC; 1.
CC Complete proteome.
CC SEQUENCE 293 AA; 32396 MW; BBDAPFC33743B4FF4 CRC64;

Query Match 50.6%; Score 44; DB 1; Length 293;
Best Local Similarity 64.3%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ASDGALVPRRLQHR 15
Db 22 ASDGCLTPEALVHR 35
|||||
22 ASDGCLTPEALVHR 35

RESULT 7
TRPH_HAEIN
ID TRPH_HAEIN STANDARD; PRT; 274 AA.
AC P44176;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein trph.
GN TRPH OR H11400.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.W.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

```


Query Match 46.68; Score 40.5; DB 1; Length 351;
Best Local Similarity 71.48; Pred. No. 10; Mismatches 1; Indels 1; Gaps 1;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 DGALVPRRLQ-HRP 16
IIIIII I I I I
DB 61 DGALVTRLLQEHQP 74

RESULT 10
OPSL_PATYE
ID OPSL_PATYE STANDARD; PRT; 499 AA.
AC QJ5973;
DT 15-JUL-1998 (Rel. 36, Created)
T 13-JUL-1998 (Rel. 36, Last sequence update)
T 30-MAY-2000 (Rel. 39, Last annotation update)
LE Rhodopsin, GQ-coupled (GQ-rhodopsin).
GN SCOP1.
OS Pateinoecten yessoensis (Ezo giant scallop) (Yesso scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
OC Pectinoidea; Pectinidae; Pateinoecten.
OX NCBI_TaxID=6573;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-EYE;
RX MEDLINE=97435252; Pubmed=9287291;
RA Kojima D., Terakita A., Ishikawa T., Tsukahara Y., Maeda A.,
RA Shichida Y.
RT "A novel Go-mediated phototransduction cascade in scallop visual
RT cells.";
RL J. Biol. Chem. 272:22979-22989(1997).
CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
CC MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
CC LINKED TO CIS-RETINAL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: RETINA. EXPRESSED IN THE DEPOLARIZING CELL
CC LAYER OF THE PHOTORECEPTOR CELLS DISTANT FROM THE LENS.
CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY
CC BE PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC OPSIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB006454; BAA22217.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN_RECEPTOR_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEPTOR_FL_2; 1.
DR PROSITE; PS00238; OPSIN; 1.
KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
KW Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.
FT DOMAIN 1 50 EXTRACELLULAR.
FT TRANSSEM 51 75 1 (POTENTIAL).
FT DOMAIN 76 87 CYTOPLASMIC.
FT TRANSSEM 88 114 2 (POTENTIAL).
FT DOMAIN 115 128 3 (POTENTIAL).
FT TRANSSEM 129 148 4 (POTENTIAL).
FT DOMAIN 149 168 5 (POTENTIAL).
FT TRANSSEM 169 192 6 (POTENTIAL).
FT DOMAIN 193 216 7 (POTENTIAL).
FT TRANSSEM 217 244 8 (POTENTIAL).
FT DOMAIN 245 278 9 (POTENTIAL).
FT TRANSSEM 279 302 10 (POTENTIAL).

QY 8 VPRRLQHRP 16
IIIIII I I I I
DB 463 IPRVQHRP 471

Query Match 46.08; Score 40; DB 1; Length 499;
Best Local Similarity 66.78; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 VPRRLQHRP 16
IIIIII I I I I
DB 463 IPRVQHRP 471

RESULT 11
SPAL_SYNY3
ID SPAL_SYNY3 STANDARD; PRT; 695 AA.
AC P74576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable arginine decarboxylase 1 (EC 4.1.1.19) (ADC).
GN SPAL OR SPEA OR SLR0662.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; Pubmed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shampo S., Takeuchi C., Wada T., Watanabe A., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: L-arginine = agmatine + CO(2).
CC -1- COPACTOR: PYRIDOXAL PHOSPHATE (By similarity).
CC -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF SPERMIDINE FROM
CC ARGININE.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90916; BAA18683.1;
DR InterPro; IPR000183; Orn_DAP_Arg_decarboxylase.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR Pfam; PF00278; Orn_DAP_Arg_dec; 2.
DR PRINTS; PR01179; ODACRBLXASE.
DR PROSITE; PS00878; ODR_DC_2.1; 1.
DR PROSITE; PS00879; ODR_DC_2.2; 1.
KW Spermidine biosynthesis; Lyase; Decarboxylase; Pyridoxal phosphate;
KW Complete proteome.
FT DOMAIN 332 342 SUBSTRATE-BINDING (POTENTIAL).
FT TRANSSEM 343 78238 MW; E90EB699D66320D CRC64;
FT DOMAIN 78238 MW; E90EB699D66320D CRC64;
FT TRANSSEM 78238 MW; E90EB699D66320D CRC64;

```

Query Match          46.0%; Score 40; DB 1; Length 695;
Best Local Similarity 72.7%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 ALVPRRLQHRP 16
DB 208 ALLARRLGHHP 218
      ||| ||| |||

RESULT 12
NIRB_ECOLI
AC P08201; STANDARD; PRT; 847 AA.
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
JE Nitrite reductase [NAD(P)H] large subunit (EC 1.6.6.4).
GN NIRB OR B3365.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=89282391; PubMed=2543955;
RA Bell A.I., Gaston K.L., Cole J.A., Busby S.J.W.;
RT "Cloning of binding sequences for the Escherichia coli transcription
RT activators, FNR and CRP: location of bases involved in discrimination
RT between FNR and CRP.";
RL Nucleic Acids Res. 17:3865-3874(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=90345936; PubMed=2200672;
RA Peakman T., Crouzet J., Mayaux J.F., Busby S.J.W., Mohan S.,
RA Harborne N., Wootton J., Nicolson R., Cole J.A.;
RT "Nucleotide sequence, organisation and structural analysis of the
RT products of genes in the nirB-cysG region of the Escherichia coli
RT K-12 chromosome.";
RL Eur. J. Biochem. 191:315-323(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=88062713; PubMed=2445993;
RA Jayaraman P.S., Peakman T.C., Busby S.J.W., Quincey R.V.,
RA Cole J.A.;
RT "Location and sequence of the promoter of the gene for the NADH-
RT dependent nitrite reductase of Escherichia coli and its regulation by
RT oxygen, the Fnr protein and nitrite.";
RL J. Mol. Biol. 196:781-788(1987).
CC -1- CATALYTIC ACTIVITY: 3 NAD(P)H + nitrite = 3 NAD(P)(+) + NH(4)OH +
CC H(2)O.
CC -1- COFACTOR: THIS ENZYME IS A FAD FLAVOPROTEIN THAT ALSO CONTAINS
CC A SIROHEME AND ONE 2FE-2S IRON-SULFUR CENTER.
CC -1- PATHWAY: SECOND STEP IN NITRATE ASSIMILATION (DENITRIFICATION).
CC -1- SUBUNIT: HOMODIMER WHICH ASSOCIATES WITH NIRD.
CC -1- SIMILARITY: THE C-TERMINAL DOMAIN IS A 4FE-4S/SIROHEME DOMAIN
CC FOUND IN NITRITE REDUCTASES (EC 1.6.6.4 AND EC 1.7.1.1) AND
CC SULFITE REDUCTASES (EC 1.8.1.2 AND EC 1.8.7.1).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC [3]

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14202; CAA32416.1; -
CC EMBL; U18997; AAA58162.1; -
CC EMBL; AE000412; AAC76390.1; -
CC PIR; S00529; S00529.
CC PIR; S10791; S10791.
CC PIR; S04349; S04349.
CC EcoGene; EG10653; nirB.
CC InterPro; IPR001327; FAD_pyr_redox.
CC InterPro; IPR000660; Nir_Sir.
CC Pfam; PF01077; Nir_Sir; 1.
CC PRINTS; PR00397; SIROHAEM.
CC PROSITE; PS00365; NIR_SIR; 1.
CC Oxidoreductase; FAD; Flavoprotein; Iron-sulfur; Nitrate assimilation;
KW Heme; NADP; Complete proteome.
FT NP_BIND 44 79 FAD (POTENTIAL).
FT METAL 193 225 NAD(P)H (POTENTIAL).
FT METAL 641 641 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 647 647 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 681 681 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 685 685 IRON-SULFUR (2FE-2S) AND SIROHEME
FT (BY SIMILARITY).
FT CONFLICT 442 442 G -> A (IN REF. 1 AND 2).
FT CONFLICT 835 847 YERIPYTLVEDNA -> MNYSQ (IN REF. 1 AND 2).
SQ SEQUENCE 847 AA; 93121 MW; 5265AD93FD390EB4 CRC64;

Query Match          46.0%; Score 40; DB 1; Length 847;
Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 LVPRRLQHRP 16
DB 822 MYPREQHRP 831
      ||| ||| |||

RESULT 13
NOS3_BOVIN
ID NOS3_BOVIN STANDARD; PRT; 1204 AA.
AC P29473;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nitric-oxide synthase, endothelial (EC 1.14.13.39) (EC-NOS) (NOS, type
DE III) (NOSIII) (Endothelial NOS) (ENOS) (Constitutive NOS) (CNOS).
GN NOS3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92335295; PubMed=1378626;
RA Lamas S., Marsden P.A., Li G.K., Tempst P., Michel T.;
RT "Endothelial nitric oxide synthase: molecular cloning and
RT characterization of a distinct constitutive enzyme isoform.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6348-6352(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93055452; PubMed=1385480;
RA Nishida K., Harrison D.G., Navas J.P., Fisher A.A., Dockery S.P.,
RA Tematsu M., Nerem R.M., Alexander R.W., Murphy T.J.;
RT "Molecular cloning and characterization of the constitutive bovine
RT aortic endothelial cell nitric oxide synthase.";
RL J. Clin. Invest. 90:2092-2096(1992).
RN [3]

```

SEQUENCE FROM N.A.
 RC TISSUE-Aortic endothelium;
 RX MEDLINE=92348367; PubMed=1379225;
 RA Sessa W.C., Harrison J.K., Barber C.M., Zeng D., Durioux M.E.,
 RA D'Angelo D.D., Lynch K.R., Peach M.J.;
 RT "Molecular cloning and expression of a cDNA encoding endothelial cell
 RT nitric oxide synthase.";
 RL J. Biol. Chem. 267:15274-15276(1992).
 RN [4]
 RP MYRISTOYLATION.
 RX MEDLINE=93231982; PubMed=7682550;
 RA Busconi L., Michel T.;
 RT "Endothelial nitric oxide synthase. N-terminal myristoylation
 RT determines subcellular localization.";
 RL J. Biol. Chem. 268:8410-8413(1993).
 RN [5]
 RP PALMITOYLATION.
 RX MEDLINE=96102197; PubMed=8524847;
 RA Robinson L.J., Michel T.;
 RT "Mutagenesis of palmitoylation sites in endothelial nitric oxide
 RT synthase identifies a novel motif for dual acylation and subcellular
 RT targeting.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:11776-11780(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 66-481.
 RX MEDLINE=99091052; PubMed=9875848;
 RA Raman C.S., Li H., Martasek P., Kral V., Masters B.S., Poulos T.L.;
 RT "Crystal structure of constitutive endothelial nitric oxide synthase:
 RT a paradigm for pterin function involving a novel metal center.";
 RL Cell 95:939-950(1998).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.86 ANGSTROMS) OF 66-481.
 RX MEDLINE=20503854; PubMed=11051558;
 RA Li H., Raman C.S., Martasek P., Kral V., Masters B.S., Poulos T.L.;
 RT "Mapping the active site polarity in structures of endothelial
 RT nitric oxide synthase heme domain complexed with isothioureas.";
 RL J. Inorg. Biochem. 81:133-139(2000).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS).
 RX MEDLINE=21229525; PubMed=11331003;
 RA Li H., Raman C.S., Martasek P., Masters B.S., Poulos T.L.;
 RT "Crystallographic studies on endothelial nitric oxide synthase
 RT complexed with nitric oxide and mechanism-based inhibitors.";
 RL Biochemistry 40:5399-5406(2001).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
 RX MEDLINE=21552923; PubMed=11695891;
 RA Raman C.S., Li H., Martasek P., Southan G., Masters B.S.,
 RA Poulos T.L.;
 RT "Crystal structure of nitric oxide synthase bound to nitro indazole
 RT reveals a novel inactivation mechanism.";
 RL Biochemistry 40:13448-13455(2001).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS).
 RX MEDLINE=21336567; PubMed=11331290;
 RA Raman C.S., Li H., Martasek P., Babu B.R., Griffith O.W., Southan G.,
 RA Masters B.S., Poulos T.L.;
 RT "Implications for isoform-selective inhibitor design derived from the
 RT binding mode of bulky isothioureas to the heme domain of endothelial
 RT nitric-oxide synthase.";
 RL J. Biol. Chem. 276:26486-26491(2001).
 CC -|- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS IMPLICATED IN
 CC VASCULAR SMOOTH MUSCLE RELAXATION THROUGH A CGMP-MEDIATED SIGNAL
 CC TRANSDUCTION PATHWAY. NO MEDIATES VASCULAR ENDOTHELIAL GROWTH
 CC FACTOR (VEGF)-INDUCED ANGIOGENESIS IN CORONARY VESSELS AND
 CC PROMOTES BLOOD CLOTTING THROUGH THE ACTIVATION OF PLATELETS.
 CC -|- CATALYTIC ACTIVITY: L-arginine + NADPH + H⁺ O₂ -> citrulline +
 CC nitric oxide + NADP(+).
 CC -|- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES
 CC TETRAHYDROBIOTIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF
 CC THE ENZYME.
 CC -|- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN.
 CC -|- SUBUNIT: HOMODIMER.

-|- SIMILARITY: BELONGS TO THE NOS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M99057; AAA30667.1; -;
 DR EMBL; M89952; AAA30494.1; -;
 DR EMBL; M95674; AAA30669.1; -;
 DR PDB; 1NSE; 18-MAY-99.
 DR PDB; 2NSE; 23-MAY-99.
 DR PDB; 3NSE; 18-MAY-99.
 DR PDB; 4NSE; 18-MAY-99.
 DR PDB; 8NSE; 21-NOV-01.
 DR PDB; 9NSE; 25-OCT-00.
 DR PDB; 1D1W; 25-OCT-00.
 DR PDB; 1ED4; 25-OCT-00.
 DR PDB; 1DM6; 13-DEC-00.
 DR PDB; 1DM7; 13-DEC-00.
 DR PDB; 1DM8; 13-DEC-00.
 DR PDB; 1DML; 20-DEC-00.
 DR PDB; 1DMJ; 20-DEC-00.
 DR PDB; 1ED3; 31-JAN-01.
 DR PDB; 1ED6; 31-JAN-01.
 DR PDB; 1FOI; 20-JUL-01.
 DR PDB; 1FOL; 20-JUL-01.
 DR PDB; 1FOO; 20-JUL-01.
 DR PDB; 1FOP; 20-JUL-01.
 DR PDB; 1DIW; 25-JUL-01.
 DR PDB; 1DI1X; 25-JUL-01.
 DR PDB; 1DI1Y; 25-JUL-01.
 DR PDB; 1D0C; 21-NOV-01.
 DR PDB; 1D00; 21-NOV-01.
 DR PDB; 1FOJ; 16-NOV-01.
 DR InterPro; IPR003097; FAD_binding.
 DR InterPro; IPR001094; Flavdn-like.
 DR InterPro; IPR001226; Flavodoxin.
 DR InterPro; IPR001709; Flavpyrid_cyt_redctse.
 DR InterPro; IPR004030; NOS.
 DR InterPro; IPR001433; Oxidored_FAD.
 DR Pfam; PF00667; FAD_binding; 1.
 DR Pfam; PF00258; Flavodoxin; 1.
 DR Pfam; PF00175; NAD_binding; 1.
 DR Pfam; PF02898; NO_synthase; 1.
 DR PRINTS; PR00369; FLAVODOXIN.
 DR PRINTS; PR00371; FPNCR.
 DR PROSITE; PS60001; NOS; 1.
 DR KX Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Myristate;
 KW Lipoprotein; Palmitate; Phosphorylation; Calcium-binding; Heme;
 KW Zinc; Metal-binding; Multigene family; 3D-structure.
 FT INIT_MET 0
 FT BINDING 185 185 HEME.
 FT DOMAIN 491 511 CALMODULIN-BINDING (POTENTIAL).
 FT NP_BIND 650 681 FMN (PYRIMIDINE PART) (BY SIMILARITY).
 FT NP_BIND 794 805 FAD (ADP PART) (BY SIMILARITY).
 FT NP_BIND 936 946 FAD (FLAVIN PART) (BY SIMILARITY).
 FT NP_BIND 1011 1029 NADP (RIBOSE PART) (BY SIMILARITY).
 FT NP_BIND 1109 1124 NADP (ADP PART) (BY SIMILARITY).
 FT LIPID 1 1 MYRISTATE.
 FT LIPID 14 14 PALMITATE.
 FT LIPID 25 25 PALMITATE.
 FT METAL 95 95 ZINC.
 FT METAL 100 100 ZINC.
 FT MOD_RES 142 142 PHOSPHORYLATION (BY CAPK).
 FT CONFLICT 99 99 C -> R (IN REF. 3).
 FT CONFLICT 164 164 Y -> I (IN REF. 3).
 FT CONFLICT 317 327 EHPLEWFAAL -> GAPHTGVVRGP (IN REF. 3).
 FT CONFLICT 454 454 S -> Y (IN REF. 3).
 FT CONFLICT 458 458 T -> P (IN REF. 3).

FT CONFLICT 740 740 T -> A (IN REF. 3).
 FT CONFLICT 803 804 CP -> SA (IN REF. 3).
 FT CONFLICT 856 856 L -> V (IN REF. 3).
 FT CONFLICT 906 907 WF -> LV (IN REF. 3).
 FT CONFLICT 1041 1041 A -> H (IN REF. 3).
 SQ SEQUENCE 1204 AA; 133155 MW; D017210062ABE4B0 CRC64;

Query Match 46.0%; Score 40; DB 1; Length 1204;
 Best Local Similarity 43.8%; Pred. No. 47;

Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 CASDGLVPRRLQHRP 16
 DB 99 CCLGSLVLPKRLQTRP 114

RESULT 14

ATPI_PORPU STANDARD; PRT; 248 AA.
 AC P51247;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE ATP synthase A chain precursor (EC 3.6.3.14) (Subunit IV).
 GN ATPI.
 OS Porphyra purpurea.
 OG Chloroplast.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
 OX NCBI_TaxID=2787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AVONPORT;
 RA Reith M.E., Munnolland J.;
 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast genome";
 RL Plant Mol. Biol. Rep. 13:333-335(1995).
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0) HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast thylakoid membrane.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U38804; AAC08133.1;
 DR Mendel; 10272; PORPU; atpi.1;
 DR InterPro; IPR000568; ATP_synt_A.
 DR Pfam; PF00119; ATP_synt_A; 1.
 DR PRINTS; PR00123; ATPASEA.
 DR PROSITE; PS00449; ATPASE_A; 1.
 KW Hydrogen ion transport; CF(0); Chloroplast; Transmembrane; Signal.
 FT SIGNAL 1 ?
 FT CHAIN 1 ?
 FT TRANSMEM 35 55
 FT TRANSMEM 94 114
 FT TRANSMEM 133 153
 FT TRANSMEM 202 222
 SQ SEQUENCE 248 AA; 27768 MW; 19BF4604B2690169 CRC64;

Query Match 44.8%; Score 39; DB 1; Length 248;
 Best Local Similarity 50.0%; Pred. No. 13;

Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 OY 1 CASDGLVPRRLQHRP 16
 DB 107 CNWAGALIPKLIHLP 122

RESULT 15

OGGI_HUMAN STANDARD; PRT; 345 AA.
 ID OGG1_HUMAN
 AC O1557; O00390; P78554; O00670; O00705; O14876; O95488; Q9UL34;
 AC Q9Y6C3; Q9Y6C4; Q9UIK0; Q9UIK1; Q9UIK2; Q9Y2C0; Q9Y2C1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE N-glycosylase/DNA lyase [Includes: 8-oxoguanine DNA glycosylase
 DE (EC 3.2.2.-); DNA-(apurinic or apyrimidinic site) lyase
 DE (EC 4.2.99.18) (AP lyase)].
 GN OGG1 OR MMH OR MUTM OR OGH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Colon;
 RX MEDLINE=97330655; PubMed=9187114;
 RA Aburatani H., Hippo Y., Ishida T., Takashima R., Matsuba C.,
 RA Kodama T., Takao M., Yasui A., Yamamoto K., Asano M., Fukasawa K.,
 RA Yoshinari T., Inoue H., Otsuka E., Nishimura S.;
 RT "Cloning and characterization of mammalian 8-hydroxyguanine-specific
 RT DNA glycosylase/apurinic, apyrimidinic lyase, a functional mutM
 RT homologue";
 RL Cancer Res. 57:2151-2156(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1A).
 RX MEDLINE=97352815; PubMed=9207108;
 RA Rosenquist T.A., Zharkov D.O., Grollman A.P.;
 RT "Cloning and characterization of a mammalian 8-oxoguanine DNA
 RT glycosylase";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:7429-7434(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97368311; PubMed=9223306;
 RA Roldan-Arjona T., Wei Y.-F., Carter K.C., Klungland A., Anselmino C.,
 RA Wang R.-P., Augustus M., Lindahl T.;
 RT "Molecular cloning and functional expression of a human cDNA encoding
 RT the antitumor enzyme 8-hydroxyguanine-DNA glycosylase";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:8016-8020(1997).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1A).
 RX MEDLINE=97368310; PubMed=9223305;
 RA Radicella J.P., Bherin C., Desmaze C., Fox M.S., Boiteux S.;
 RT "Cloning and characterization of hOGG1, a human homolog of the OGG1
 RT gene of Saccharomyces cerevisiae";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:8010-8015(1997).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1A).
 RX MEDLINE=97342862; PubMed=9197244;
 RA Lu R., Nash H.M., Verdine G.L.;
 RT "A mammalian DNA repair enzyme that excises oxidatively damaged
 RT guanines maps to a locus frequently lost in lung cancer";
 RL Curr. Biol. 7:397-407(1997).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1A).
 RX MEDLINE=97334205; PubMed=9190902;
 RA Arai K., Morishita K., Shimura K., Kohno T., Taniwaki M., Ohwada S.,
 RA Yokota J.;
 RT "Cloning of a human homolog of the yeast OGG1 gene that is involved in
 RT the repair of oxidative DNA damage";
 RL Oncogene 14:2857-2861(1997).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 1A).

```

RX MEDLINE=98012228; PubMed=9348312;
RA Kuo F.C., Sklar J.L.;
RT "Augmented expression of a human gene for 8-oxoguanine DNA glycosylase
RT (MUGM) in B lymphocytes of the dark zone in lymph node germinal
RT centers.";
RL J. Exp. Med. 186:1547-1556(1997).
RN [8]
RP SEQUENCE FROM N.A. (ISOFORM 1A).
RX MEDLINE=98026907; PubMed=9321410;
RA Bjoras M., Luna L., Johnson B.E., Hoff E., Haug T., Rognes T.,
RA Seeberg E.;
RT "Opposite base-dependent reactions of a human base excision repair
RT enzyme on DNA containing 7,8-dihydro-8-oxoguanine and abasic sites.";
RL EMBO J. 16:6314-6322(1997).
RN [9]
RP SEQUENCE FROM N.A. (ISOFORM 1A).
RA Dhenaut A., Boiteux S., Radicella J.;
RT "Genomic structure and promoter characterization of the human 8-OH-
RT guanine glycosylase gene (OGG1) gene.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [10]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99380087; PubMed=10449904;
RA Ishida T., Hippo Y., Nakahori Y., Matsushita I., Kodama T.,
RA Nishimura S., Aburatani H.;
RT "Structure and chromosome location of human OGG1.";
RL Cytogenet. Cell Genet. 85:232-236(1999).
RN [11]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99250167; PubMed=10233168;
RA Nishioka K., Ohtsubo T., Oda H., Fujiwara T., Kang D., Sugimachi K.,
RA Nakabeppu Y.;
RT "Expression and differential intracellular localization of two major
RT forms of human 8-oxoguanine DNA glycosylase encoded by alternatively
RT spliced OGG1 mRNAs.";
RL Mol. Biol. Cell 10:1637-1652(1999).
RN [12]
RP REVIEW.
RX MEDLINE=20239648; PubMed=10775435;
RA Boiteux S., Radicella J.P.;
RT "The human OGG1 gene: structure, functions, and its implication in the
RT process of carcinogenesis.";
RL Arch. Biochem. Biophys. 377:1-8(2000).
RN [13]
RP VARIANT HIS-154.
RX MEDLINE=98438755; PubMed=9765618;
RA Shimura K., Kohno T., Kasai H., Koda K., Sugimura H., Yokota J.;
RT "Infrequent mutations of the hOGG1 gene, that is involved in the
RT excision of 8-hydroxyguanine in damaged DNA, in human gastric
RT cancer.";
RL Jpn. J. Cancer Res. 89:825-828(1998).
RN [14]
RP VARIANT GLN-46.
RX MEDLINE=98324718; PubMed=9662341;
RA Chevillard S., Radicella J.P., Levalois C., Lebeau J., Poupon M.F.,
RA Oudard S., Dutrillaux B., Boiteux S.;
RT "Mutations in OGG1, a gene involved in the repair of oxidative DNA
RT damage, are found in human lung and kidney tumours.";
RL Oncogene 15:3083-3086(1998).
RN [15]
RP CHARACTERIZATION OF VARIANT CYS-326.
RX MEDLINE=99428653; PubMed=10497264;
RA Dherin C., Radicella J.P., Dizdareglu M., Boiteux S.;
RT "Excision of oxidatively damaged DNA bases by the human alpha-hOgg1
RT protein and the polymorphic alpha-hOgg1(Ser326Cys) protein which is
RT frequently found in human populations.";
RL Nucleic Acids Res. 27:4001-4007(1999).
RN [16]
RP CHARACTERIZATION OF VARIANTS GLN-46; HIS-154 AND CYS-326.
RX MEDLINE=20368626; PubMed=10908322;
RA Audebert M., Radicella J.P., Dizdareglu M.;
RT "Effect of single mutations in the OGG1 gene found in human tumors on
RT the substrate specificity of the ogg1 protein.";

```

```

RL Nucleic Acids Res. 28:2672-2678(2000).
CC -!- FUNCTION: DNA REPAIR ENZYME THAT INCISES DNA AT 8-OXOG RESIDUES.
CC EXCISES 7,8-DIHYDRO-8-OXOGUANINE AND 2,6-DIAMINO-4-HYDROXY-5-N-
CC METHYLFORMAMIDOPYRIMIDINE (FAPY) FROM DAMAGED DNA. HAS A BETA-
CC LASE ACTIVITY THAT NICKS DNA 3' TO THE LESION.
CC -!- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE NEAR APURINIC OR
CC APYRIMIDINIC SITES TO PRODUCTS WITH 5'-PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (ISOFORM 1A) AND MITOCHONDRIAL
CC (ISOFORM 2A).
CC -!- ALTERNATIVE PRODUCTS: 8 ISOFORMS; 1A/ALPHA (SHOWN HERE), 1B, 1C,
CC 2A/BETA, 2B, 2C, 2D AND 2E; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC 1A IS THE PREVALENT FORM.
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
CC -!- DISEASE: DEFECTS IN OGG1 ARE ASSOCIATED WITH TUMOR FORMATION.
CC -!- SIMILARITY: BELONGS TO THE OGG1 FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U96710; AAB81132.1; -
CC EMBL; Y11838; CAA72536.1; -
CC EMBL; Y11731; CAA72414.1; -
CC EMBL; AB000410; BAA19103.1; -
CC EMBL; AF003595; AAB61340.1; -
CC EMBL; U88527; AAB68614.1; -
CC EMBL; U88620; AAB68615.1; -
CC EMBL; Y13277; CAA73726.1; -
CC EMBL; AF026691; AAB84013.1; -
CC EMBL; AJ131341; CAA10351.1; -
CC EMBL; AF088282; AAD41680.1; -
CC EMBL; AF088282; AAD41681.1; -
CC EMBL; AF088282; AAD41682.1; -
CC EMBL; AB019528; BAA76635.1; -
CC EMBL; AB019529; BAA76636.1; -
CC EMBL; AB019530; BAA76637.1; -
CC EMBL; AB019531; BAA76638.1; -
CC EMBL; AB019532; BAA76639.1; -
CC MIM; 601982; -
CC InterPro: IPR003265; Endo_3c.
CC Pfam; PF00730; HhH-GPD; 1.
CC KW Hydrolase; Nuclease; Endonuclease; Lyase; DNA repair; Glycosidase;
CC Multifunctional enzyme; Nuclear protein; Mitochondrion;
CC Alternative splicing; Polymorphism.
CC ACT_SITE 249 249 BY SIMILARITY.
CC VARSPPLIC 317 345 VLFSADLRQSRHQAQEPAPKRRKSGKSGPEG -> VSVPRCP
CC (IN ISOFORM 1B).
CC VARSPPLIC 317 345 VLFSADLRQSRHQAQEPAPKRRKSGKSGPEG -> TTPSYRCC
CC SVPTCANPAMLRSHQOASAEVRPKGRKARWGLDKEIFQAPS
CC PPFTSLSPSPSPSLMLGRGLPVTTSKARHPQIKQSVCTRW
CC GGGY (IN ISOFORM 1C).
CC VARSPPLIC 317 345 VLFSADLRQSRHQAQEPAPKRRKSGKSGPEG -> GLLGNAFD
CC GHQLRLPLIFCQDLHREGPPPIGRGDSOGEELEPQLPSLSS
CC IPYFCDHCWTKDVPDPLVTHPSGSRDGHMTQAMPVKVY
CC SPLATVIGHYMQASLLAL (IN ISOFORM 2A).
CC VARSPPLIC 250 345 VADICICLMALDKDQPAVPDVMHIAQORDYSWHPTTSQAKG
CC PSPQTKELNGFERSLWGPVAGAAQVLFSAADLRQSRHQAQ
CC PPAKRRKSGKSGPEG -> GLLGNAFDGHQLRLPLIFCQDLH
CC REGPPPIGRGDSOGEELEPQLPSLSSIPYFCDHCWTKD
CC -----
CC Query Match 44.8%; Score 39; DB 1; Length 345;
CC Best Local Similarity 70.0%; Pred. No. 19;
CC Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CC -----
CC QY 6 ALVPRRLQHR 15
CC ||:||||: ||
CC Db 5 ALLPRRMGHR 14

```

Search completed: September 10, 2002, 10:46:42
Job time: 120 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2002, 10:43:37 ; Search time 15.15 Seconds
(without alignments)
101.480 Million cell updates/sec

Title: US-09-674-254-3

Perfect score: 87

Sequence: 1 CASDGLVPRRLQHRP 16

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	75.9	368	2 S67507	morphogen lefty pr
2	45	51.7	293	2 E64874	probable metal-dep
3	45	51.7	293	2 F90858	probable enzymes
4	45	51.7	293	2 A85761	probable enzymes y
5	44	50.6	293	2 AG0653	conserved hypotet
6	44	50.6	522	2 D81900	hypothetical prote
7	44	50.6	522	2 A81124	conserved hypotet
8	43	49.4	274	2 F64027	hypothetical prote
9	42.5	48.9	656	2 S55262	UV-endonuclease
10	42	48.3	335	2 D38532	hypB protein - Rho
11	41.5	47.7	327	1 S72913	cytochrome c-type
12	41	47.1	606	2 A01219	diol dehydratase-r
13	41	47.1	606	2 AG1572	hypB homolog - Xan
14	40.5	46.6	351	2 B49906	germin homolog Ger
15	40	46.0	106	2 T06563	dihydropteroate sy
16	40	46.0	335	2 H75511	hypothetical prote
17	40	46.0	289	2 F84411	arginine decarboxy
18	40	46.0	695	2 S76771	nitrite reductase
19	40	46.0	847	2 H31155	nitrite reductase
20	40	46.0	847	2 H65130	nitrite reductase
21	40	46.0	847	2 E86001	nitrite reductase
22	40	46.0	1205	1 A38943	nitric-oxide synth
23	39.5	45.4	262	2 E98265	hypothetical prote
24	39.5	45.4	262	2 A03019	conserved hypotet
25	39	44.8	109	2 T49452	hypothetical prote
26	39	44.8	248	2 S73168	H+-transporting AT
27	39	44.8	346	2 T45069	8-hydroxy-guanine
28	39	44.8	545	2 A87448	conserved hypotet
29	39	44.8	682	2 T28899	hypothetical prote

ALIGNMENTS

RESULT 1

S67507

morphogen lefty precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999

C:Accession: S67507

R:Meno, C.; Saijoh, Y.; Fujii, H.; Ikeda, M.; Yokoyama, M.; Toyoda, Y.;

Nature 381, 151-155, 1996

A:Title: Left-right asymmetric expression of the TGF-beta-family member lefty in mous

A:Reference number: S67507; MUID:96202359

A:Accession: S67507

A:Molecule type: mRNA

A:Residues: 1-368 <MEN>

A:Cross-references: EMBL:D83921; NID:gl325920; PIDN:BAAL2121.1; PID:d1012795; PID:gl4

A:Note: the authors translated the codon ACG for residue 241 as His

C:Keywords: growth factor

F:78-368/Product: morphogen lefty #status predicted <MAT1>

F:136-368/Product: morphogen lefty #status predicted <MAT2>

Query Match

Best Local Similarity 75.9%; Score 66; DB 2; Length 368;

Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQ 13

Db 355 CASDGLVPRRLQ 367

RESULT 2

E64874

probable metal-dependent phosphoesterase yciw - Escherichia coli

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 29-Sep-1999

C:Accession: E64874

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.: Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: E64874

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-293 <LAT>

A:Cross-references: GB:AE000224; GB:U00096; NID:gl787509; PIDN:AAC74348.1; PID:gl7875

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: yciw

C:Superfamily: hypothetical protein HI1400

Query Match

51.7%; Score 45; DB 2; Length 293;

Best Local Similarity 64.3%; Pred. No. 3.3; Mismatches 0; Gaps 0;

Matches 9; Conservative 0; Indels 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHR 15
 ||||| | | | |
 Db 18 ASDGCLTPALVHR 31

RESULT 3
 F90858
 probable enzymes [imported] - Escherichia coli (strain O157:H7, substrain R1MD 0509952)
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: F90858
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.;
 Jaisawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 JNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11250796
 A:Accession: F90858
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-293 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA035261.1; PID:g13361303; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain R1MD 0509952
 A:Genetics:
 A:Gene: ECs1838
 A:Superfamily: hypothetical protein H11400

Query Match 51.7%; Score 45; DB 2; Length 293;
 Best Local Similarity 64.3%; Pred. No. 3.3;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHR 15
 ||||| | | | |
 Db 18 ASDGCLTPALVHR 31

RESULT 4
 A85761
 probable enzymes yciI yciV [imported] - Escherichia coli (strain O157:H7, substrain EDL9
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: A85761
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: A85761
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-293 <STO>
 A:Cross-references: GB:AE005174; NID:g12515531; PIDN:AA056549.1; GSPDB:GN00145; UWGP:225
 A:Experimental source: strain O157:H7, substrain EDL933
 A:Genetics:
 A:Gene: yciV
 A:Superfamily: hypothetical protein H11400

Query Match 51.7%; Score 45; DB 2; Length 293;
 Best Local Similarity 64.3%; Pred. No. 3.3;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHR 15
 ||||| | | | |
 Db 18 ASDGCLTPALVHR 31

RESULT 5
 GC0653
 conserved hypothetical protein STY1329 [imported] - Salmonella enterica subsp. enterica

C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AG0653
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0653
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-293 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08410.1; PID:g16502453; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1329
 C:Superfamily: hypothetical protein H11400

Query Match 50.6%; Score 44; DB 2; Length 293;
 Best Local Similarity 64.3%; Pred. No. 5;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHR 15
 ||||| | | | |
 Db 22 ASDGCLTPETLVR 35

RESULT 6
 DB1900
 hypothetical protein NMA1314 [imported] - Neisseria meningitidis (strain 22491 serogr
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: DB1900
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
 A:Reference number: A81775; MUID:20222556
 A:Accession: DB1900
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-522 <PAR>
 A:Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84566.1; PID:g737
 A:Experimental source: serogroup A, strain 22491
 C:Genetics:
 A:Gene: NMA1314
 C:Superfamily: Haemophilus influenzae hypothetical protein HI1501

Query Match 50.6%; Score 44; DB 2; Length 522;
 Best Local Similarity 57.1%; Pred. No. 8.9;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 SDGALVPRRLQHRP 16
 ||||| | | | |
 Db 143 SDGLYLPFRFIHRP 156

RESULT 7
 A81124
 conserved hypothetical protein NMB1095 [imported] - Neisseria meningitidis (strain MC
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: A81124
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
 xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755
 A:Accession: A81124
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-522 <TET>
 A:Cross-references: GB:AE002459; GB:AE002098; NID:g7226320; PIDN:AAF41487.1; PID:g7226320
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NM81095
 C:Superfamily: Haemophilus influenzae hypothetical protein H11501

Query Match 50.6%; Score 44; DB 2; Length 522;
 Best Local Similarity 57.1%; Pred. No. 8.9;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Y 3 SDGALVPRRLQHRP 16
 ||| :|| :|||
 Db 143 SDGLYPRNFTHRP 156

RESULT 8
 F64027
 hypothetical protein H11400 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
 C:Accession: F64027
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervilave, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: F64027
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-274 <TIGR>
 A:Cross-references: GB:U32820; GB:I42023; NID:gi574231; PIDN:AAC23046.1; PID:gi574235; PID:gi574235; PID:gi574235
 C:Superfamily: hypothetical protein H11400

Query Match 49.4%; Score 43; DB 2; Length 274;
 Best Local Similarity 64.3%; Pred. No. 7;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHR 15
 ||||| :|||
 Db 13 ASDGVLSPTELVHR 26

RESULT 9
 S55262
 UV-endonuclease - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
 C:Accession: S55262
 R:Yajima, H.; Takao, M.; Yasuhira, S.; Zhao, J.H.; Ishii, C.; Inoue, H.; Yasui, A. EMBO J. 14, 2393-2399, 1995
 A:Title: A eukaryotic gene encoding an endonuclease that specifically repairs DNA damage
 A:Reference number: S55262; MUID:9522980
 A:Accession: S55262
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-656 <YAJ>
 A:Cross-references: EMBL:D11392; NID:gi526560; PID:g927215
 C:Genetics:
 A:Gene: uvel

Query Match 48.9%; Score 42.5; DB 2; Length 656;
 Best Local Similarity 64.3%; Pred. No. 21;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 4 DGALVPR-RLQHRP 16
 ||| :|| :|||
 Db 495 DGAVTPRRRKHHP 508

RESULT 10
 D38532
 hypB protein - Rhodobacter capsulatus
 C:Species: Rhodobacter capsulatus
 C:Date: 12-Jul-1991 #sequence_revision 12-Jul-1991 #text_change 24-Sep-1999
 C:Accession: D38532; S32950; S21903
 R:Xu, H.W.; Wall, J.D.
 J. Bacteriol. 173, 2401-2405, 1991
 A:Title: Clustering of genes necessary for hydrogen oxidation in Rhodobacter capsulatus
 A:Reference number: A38532; MUID:91177833
 A:Accession: D38532
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-335 <XUA>
 A:Cross-references: GB:M55089; NID:gi151949; PIDN:AAA72926.1; PID:gi151953; EMBL:X61007
 R:Colbeau, A.; Richaud, P.; Toussaint, B.; Caballero, F.J.; Elster, C.; Delphin, C.; Mol. Microbiol. 8, 15-29, 1993
 A:Title: Organization of the genes necessary for hydrogenase expression in Rhodobacter
 A:Reference number: S32941; MUID:93268090
 A:Accession: S32950
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-335 <COL>
 A:Cross-references: EMBL:X61007; NID:g46051; PIDN:CAA43326.1; PID:g46053
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 C:Genetics:
 A:Gene: hypB
 C:Superfamily: hydrogenase expression/formaton protein hypB

Query Match 48.3%; Score 42; DB 2; Length 335;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQHRP 16
 ||||| :|||
 Db 208 CHLDGAMVEQALHLP 223

RESULT 11
 S72913
 cytochrome c-type synthesis protein homolog - Mycobacterium leprae
 N:Alternate names: hypothetical protein B2168_C3_281
 C:Species: Mycobacterium leprae
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 23-Mar-2001
 C:Accession: S72913
 R:Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, November 1993
 A:Description: Mycobacterium leprae cosmid B2168.
 A:Reference number: S72586
 A:Accession: S72913
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-327 <SMT>
 A:Cross-references: EMBL:U00018; NID:g467037; PIDN:AAA17249.1; PID:g467065
 C:Superfamily: cytochrome c-type synthesis protein

Query Match 47.7%; Score 41.5; DB 1; Length 327;
 Best Local Similarity 52.6%; Pred. No. 15;
 Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 1 CASD---GALVPRRLQHRP 16
 ||| :||| :|||
 Db 119 CFSELYAGALVLRRLQYRP 137

RESULT 12
AD1219
diol dehydratase-reactivating factor large chain homolog lmoll56 [Imported] - Listeria m
Species: Listeria monocytogenes
Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
Accession: AD1219
Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ck, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
Reference number: AB1077; MUID:21537279; PMID:11679669
Accession: AD1219
Status: preliminary
Molecule type: DNA
Residues: 1-606 <GLA>
Cross-references: GB:NC_003210; PIDN:CAC99234.1; PID:g16410572; GSPDB:GN00177
Experimental source: strain EGD-e
Genetics:
Gene: lmoll56
Superfamily: Klebsiella oxytoca diol dehydratase-reactivating factor large chain

Query Match 47.1%; Score 41; DB 2; Length 606;
Best Local Similarity 57.1%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 2 ASDGALVPRRLQHR 15
| | | | | | | | | |
Db 165 ADDGVLVHNRNLHK 178

RESULT 13
AG1572
diol dehydratase-reactivating factor large chain homolog linl120 [Imported] - Listeria i
Species: Listeria innocua
Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
Accession: AG1572
Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ck, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative genomics of Listeria species.
Reference number: AB1077; MUID:21537279; PMID:11679669
Accession: AG1572
Status: preliminary
Molecule type: DNA
Residues: 1-606 <GLA>
Cross-references: GB:AL592022; PIDN:CAC96351.1; PID:g164113579; GSPDB:GN00178
Experimental source: strain C11p11262
Genetics:
Gene: linl120
Superfamily: Klebsiella oxytoca diol dehydratase-reactivating factor large chain

Query Match 47.1%; Score 41; DB 2; Length 606;
Best Local Similarity 57.1%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 2 ASDGALVPRRLQHR 15
| | | | | | | | | |
Db 165 ADDGVLVHNRNLHK 178

RESULT 14
I49906
fbB homolog - Xanthomonas campestris pv. campestris
Species: Xanthomonas campestris pv. campestris
Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 16-Jul-1999

C:Accession: B49906
R:Koeplin, R.; Wang, G.; Hoette, B.; Priefer, U.B.; Puehler, A.
J. Bacteriol. 175, 7786-7792, 1993
A:Title: A 3.9-kb DNA region of Xanthomonas campestris pv. campestris that is necessary
A:Reference number: A49906; MUID:94075213
Accession: B49906
Status: preliminary
Molecule type: DNA
Residues: 1-351 <KOE>
Cross-references: GB:L23941; MID:g398116; PIDN:AAA16192.1; PID:g398120
Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homolo
F:4-334/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 46.6%; Score 40.5; DB 2; Length 351;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
Qy 4 DGALVPRRLQ-HRP 16
| | | | | | | | | |
Db 61 DGALVTRLLQEHQ 74

RESULT 15
T06563
germin homolog Ger4 - wheat (fragment)
Species: Triticum aestivum (common wheat)
Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
Accession: T06563
Bashardes, S.
submitted to the EMBL Data Library, December 1996
Reference number: 215759
Accession: T06563
Status: preliminary; translated from GB/EMBL/DDBJ
Molecule type: mRNA
Residues: 1-106 <BAS>
Cross-references: EMBL:Y09918; PIDN:CAA71053.1
Experimental source: germinating embryos
Genetics:
Gene: Ger4
Superfamily: germin
Keywords: cell wall

Query Match 46.0%; Score 40; DB 2; Length 106;
Best Local Similarity 50.0%; Pred. No. 9.1;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CASDGAIVPRRLQ 14
| | | | | | | | | |
Db 29 CAGETFLIPRGLMH 42

Search completed: September 10, 2002, 10:45:54
Job time: 137 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2002, 10:40:47 ; Search time 30.11 Seconds
(without alignments)
59.023-Million cell updates/sec

Title: US-09-674-254-3

Perfect score: 87

Sequence: 1 CASDGLVPRRLQHRP 16

-Scoring table: BLOSUM62

Gapop 10.0 , Gapext 10.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	100.0	16	21	AA1980
2	87	100.0	294	22	AA1981
3	87	100.0	370	21	AA1982
4	67	77.0	120	20	AA1983
5	67	77.0	120	20	AA1984
6	67	77.0	366	20	AA1985
7	67	77.0	366	20	AA1986
8	67	77.0	366	20	AA1987
9	67	77.0	366	20	AA1988
10	67	77.0	366	20	AA1989
11	67	77.0	366	21	AA1990

12	67	77.0	366	22	AA1991	Human protein seq
13	67	77.0	366	22	AA1992	PRO317 Homo sapi
14	67	77.0	366	22	AA1993	Human PRO317 prote
15	67	77.0	366	22	AA1994	Endometrial bleed
16	66	75.9	368	20	AA1995	Murine lefty prote
17	47	54.0	380	21	AA1996	Neisseria meningit
18	45	51.7	1643	22	AA1997	C glutamicum prote
19	44	50.6	67	20	AA1998	Extended human seq
20	44	50.6	67	21	AA1999	UCK-3 amino acid s
21	44	50.6	77	19	AA2000	Human secreted pro
22	44	50.6	77	20	AA2001	Human secreted pro
23	44	50.6	77	22	AA2002	Human EST encoded
24	44	50.6	77	22	AA2003	Human EST encoded
25	44	50.6	78	19	AA2004	Human secreted pro
26	44	50.6	90	21	AA2005	Human secreted pro
27	44	50.6	99	20	AA2006	Human secreted pro
28	44	50.6	99	21	AA2007	Human secreted pro
29	44	50.6	99	21	AA2008	UCK-1 amino acid s
30	44	50.6	99	22	AA2009	Human protein clon
31	44	50.6	99	22	AA2010	Novel human secret
32	44	50.6	102	16	AA2011	Human secreted pro
33	44	50.6	102	21	AA2012	Hepatitis GB virus
34	44	50.6	113	22	AA2013	Hepatitis GB virus
35	44	50.6	120	21	AA2014	Human colon cancer
36	44	50.6	120	21	AA2015	UCK-4 amino acid s
37	44	50.6	152	20	AA2016	Human PRO772 prote
38	44	50.6	152	21	AA2017	Human PRO772 (UNO
39	44	50.6	152	21	AA2018	Human secreted pro
40	44	50.6	152	21	AA2019	Human secreted pro
41	44	50.6	152	21	AA2020	UCK-2 amino acid s
42	44	50.6	152	21	AA2021	Human transmembran
43	44	50.6	152	22	AA2022	Novel human secret
44	44	50.6	152	22	AA2023	Human secreted pro
45	44	50.6	264	16	AA2024	Hepatitis GB virus

ALIGNMENTS

RESULT 1
AA1980
ID AA198026 standard; peptide: 16 AA.
XX AA198026;
XX AC
XX AC
XX DT
XX DT
XX DE
XX DE
XX KW
XX KW
XX KW
XX KW
XX OS
XX OS
XX PN
XX PN
XX PD
XX PD
XX PF
XX PF
XX PR
XX PR
XX PA
XX PA
XX PI
XX PI
XX DR
XX DR
XX PT
XX PT
XX PT
XX PT

Human ebaf protein antigenic C-terminal peptide.

Antigenic; ebaf; human; fertility; diagnosis; endometrial irregularity; endometrial bleeding-associated factor; contraceptive; endometrium; fertility-regulating; uterine bleeding; non-receptiveness; marker; infertility; conception; in vitro fertilization; artificial insemination

Homo sapiens.

WO9955902-A1.

04-NOV-1999.

29-APR-1999; 99WO-US09366.

29-APR-1998; 98US-0083418.

(UYSF-) UNIV SOUTH FLORIDA.

Tabibzadeh S;

WPI; 2000-052811/04.

Diagnosis of endometrial irregularities by detecting ebaf or its splice variants, particularly for diagnosing infertility

PS Claim 39; Page 95; 112pp; English.

CC This invention describes a novel method for diagnosing endometrial
CC irregularities by screening an endometrial sample, or body fluid, for
CC the presence of ebaf (endometrial bleeding-associated factor) or its
CC splice variants that are differentially expressed. The product of the
CC invention has contraceptive and fertility-regulating activity. Ebaf is
CC associated with natural or abnormal uterine bleeding and is a marker for
CC non-receptiveness of the endometrium, and is aberrantly expressed in a
CC subset of infertile women. The method is used for diagnosis, monitoring
CC and prognosis of infertility, endometriosis and abnormal uterine
CC bleeding, particularly to determine endometrial receptivity (i.e. for
CC optimizing the time of conception, including by in vitro fertilization
CC and artificial insemination methods). Ebaf is also useful as a
CC contraceptive while agents that down regulate it, e.g. antisense
CC sequences, are used to treat endometrial disorders, specifically to
CC restore fertility. This sequence represents an antigenic peptide derived
CC from the human ebaf protein C-terminal.

XX Sequence 16 AA;

Query Match 100.0%; Score 87; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 7e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AY 1 CASDGLVPRRLQHRP 16
| | | | | | | | | | | | | | | |
b 1 casdgalvprlqhrp 16

RESULT 2
AAB35940
D AAB35940 standard; protein; 294 AA.

XX AAB35940;
XX 26-FEB-2001 (first entry)

XX TGF-beta 4 amino acid sequence.
XX Heparin binding; vascular graft; matrix; cell adhesion; growth factor;
XX wound healing; dermal wound; wound healing; TGF-beta 4.
XX Unidentified.

XX WO200064481-A1.
XX 02-NOV-2000.

XX 22-APR-1999; 99WO-IB00800.
XX 22-APR-1999; 99WO-IB00800.

XX (ETHZ-) ETH ZURICH & UNIV ZURICH.

XX Sakiyama SE, Hubbell JA;
XX WPI; 2001-024627/03.

XX Matrix for controlled release of growth factor for wound healing, has
XX substrate that attaches heparin binding peptide, protein growth factor
XX that bind heparin with low affinity, and heparin or heparin-like
XX polymer

XX Example 5; Page 21; 48pp; English.

XX This invention relates to a matrix comprising a substrate capable of
XX providing attachment of a heparin binding peptide (HBP), a peptide
XX comprising a binding domain which binds heparin with high affinity,
XX heparin or heparin-like polymer, and a protein growth factor or peptide
XX fragment which has a domain that binds heparin with low affinity.
XX Included in the invention is a vascular graft comprising the matrix,

CC which is capable of supporting cell adhesion. The matrix is used for
CC delivering low heparin binding affinity growth factor proteins or
CC peptides in a controlled manner suitable for wound healing. The matrix
CC can be used in an article for treating dermal wounds, and in an
CC implantable sterilized composition capable of supporting cell adhesion.
CC The present sequence represents a growth factor protein. The protein is
CC used in an example illustrating that non-heparin-binding growth factors
CC can be released in a controlled manner from heparin-based drug delivery
CC systems based on their low affinity for heparin.

XX Sequence 294 AA;

Query Match 100.0%; Score 87; DB 22; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQHRP 16
| | | | | | | | | | | | | | | |
Db 275 casdgalvprlqhrp 290

RESULT 3
AAY92013
ID AAY92013 standard; Protein; 370 AA.

XX AAY92013;
XX 19-JUL-2000 (first entry)

XX Human transforming growth factor beta 4/ebaf monomer.

XX human transforming growth factor beta 4 monomer; ebaf; CKGF; mutant;
XX cysteine knot growth factor; hairpin loop; infertility.

XX Homo sapiens.

XX Key Location/Qualifiers
XX Misc-difference 1..266
XX /note= "optionally mutated to increase electrostatic
XX interaction between beta hairpin structure and
XX a receptor"

XX Domain 267..287
XX /label= beta_hairpin_loop_1
XX /note= "mutant optionally comprises one or more
XX substitutions in these residues"

XX Misc-difference 288..317
XX /note= "optionally mutated to increase electrostatic
XX interaction between beta hairpin structure and
XX a receptor"

XX Domain 318..337
XX /label= beta_hairpin_loop_3
XX /note= "mutant optionally comprises one or more
XX substitutions in these residues"

XX Misc-difference 338..370
XX /note= "optionally mutated to increase electrostatic
XX interaction between beta hairpin structure and
XX a receptor"

XX WO200017360-A1.

XX 30-MAR-2000.

XX 19-MAR-1999; 99WO-US05908.

XX 22-SEP-1998; 98WO-US19772.

XX (DYMA-) UNIV MARYLAND BALTIMORE.

XX Weintraub BD, Szkudlinski MW;

XX WPI; 2000-283585/24.

1	OTHER INFORMATION:	/note= "Potential proline directed
2	OTHER INFORMATION:	phosphorylation site"
3	FEATURE:	
4	NAME/KEY: Domain	
5	LOCATION: 97..98	
6	OTHER INFORMATION:	/note= "Potential proline directed
7	OTHER INFORMATION:	phosphorylation site"
8	FEATURE:	
9	NAME/KEY: Domain	
10	LOCATION: 116..117	
11	OTHER INFORMATION:	/note= "Potential proline directed
12	OTHER INFORMATION:	phosphorylation site"
13	FEATURE:	
14	NAME/KEY: Domain	
15	LOCATION: 282..283	
16	OTHER INFORMATION:	/note= "Potential proline
17	OTHER INFORMATION:	phosphorylation site"
18	FEATURE:	
19	NAME/KEY: Domain	
20	LOCATION: 459..460	
21	OTHER INFORMATION:	/note= "Potential proline directed
22	OTHER INFORMATION:	phosphorylation site"
23	FEATURE:	
24	NAME/KEY: Domain	
25	LOCATION: 472..473	
26	OTHER INFORMATION:	/note= "Potential proline directed
27	OTHER INFORMATION:	phosphorylation site"
28	FEATURE:	
29	NAME/KEY: Domain	
30	LOCATION: 602..603	
31	OTHER INFORMATION:	/note= "Potential proline directed
32	OTHER INFORMATION:	phosphorylation site"
33	FEATURE:	
34	NAME/KEY: Domain	
35	LOCATION: 727..728	
36	OTHER INFORMATION:	/note= "Potential proline directed
37	OTHER INFORMATION:	phosphorylation site"
38	FEATURE:	
39	NAME/KEY: Domain	
40	LOCATION: 838..839	
41	OTHER INFORMATION:	/note= "Potential proline directed
42	OTHER INFORMATION:	phosphorylation site"
43	FEATURE:	
44	NAME/KEY: Domain	
45	LOCATION: 869..870	
46	OTHER INFORMATION:	/note= "Potential proline directed
47	OTHER INFORMATION:	phosphorylation site"
48	FEATURE:	
49	NAME/KEY: Domain	
50	LOCATION: 872..873	
51	OTHER INFORMATION:	/note= "Potential proline directed
52	OTHER INFORMATION:	phosphorylation site"
53	FEATURE:	
54	NAME/KEY: Domain	
55	LOCATION: 1085..1086	
56	OTHER INFORMATION:	/note= "Potential proline directed
57	OTHER INFORMATION:	phosphorylation site"
58	FEATURE:	
59	NAME/KEY: Domain	
60	LOCATION: 1202..1203	
61	OTHER INFORMATION:	/note= "Potential proline directed
62	OTHER INFORMATION:	phosphorylation site"
63	FEATURE:	
64	NAME/KEY: Domain	
65	LOCATION: 114..116	
66	OTHER INFORMATION:	/note= "cAMP dependent
67	OTHER INFORMATION:	phosphorylation site"
68	FEATURE:	
69	NAME/KEY: Domain	
70	LOCATION: 141..143	
71	OTHER INFORMATION:	/note= "cAMP dependent
72	OTHER INFORMATION:	phosphorylation site"
73	FEATURE:	

```

1 NAME/KEY: Domain
2 LOCATION: 168..170
3 OTHER INFORMATION: /note= "CAMP dependent
4 OTHER INFORMATION: phosphorylation site"
5 FEATURE:
6 NAME/KEY: Domain
7 LOCATION: 633..635
8 OTHER INFORMATION: /note= "CAMP dependent
9 OTHER INFORMATION: phosphorylation site"
10 FEATURE:
11 NAME/KEY: Domain
12 LOCATION: 836..838
13 OTHER INFORMATION: /note= "CAMP dependent
14 OTHER INFORMATION: phosphorylation site"
15 FEATURE:
16 NAME/KEY: Domain
17 LOCATION: 1051..1053
18 OTHER INFORMATION: /note= "CAMP dependent
19 OTHER INFORMATION: phosphorylation site"
20 FEATURE:
21 NAME/KEY: Domain
22 LOCATION: 738..740
23 OTHER INFORMATION: /note= "CAMP dependent
24 OTHER INFORMATION: phosphorylation site"
25 US-07-908-245-2

```

Query Match 46.0%; Score 40; DB 1; Length 1205;
Best Local Similarity 43.8%; Pred. No. 66;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CASD GALVPRRLQHRP 16
| : : : : :
Db 100 CCLGSLVLP RKLOTRP 115

```

RESULT      4
US-09-428-517-3
: Sequence 3, Application US/09428517
: Patent No. 6251636
: GENERAL INFORMATION:
: APPLICANT: Betlach, Mary C.
: APPLICANT: Shah, Sanjay Krishnakant
: APPLICANT: McDaniel, Robert
: APPLICANT: Tang, Li
: TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
: FILE REFERENCE: 30062-20029.00
: CURRENT APPLICATION NUMBER: US/09/428,517
: CURRENT FILING DATE: 1999-10-28
: EARLIER APPLICATION NUMBER: 60/120,254
: EARLIER FILING DATE: 1999-02-16
: EARLIER APPLICATION NUMBER: 60/106,100
: EARLIER FILING DATE: 1998-10-29
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3
: LENGTH: 3816
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Recombinant
: OTHER INFORMATION: Oleandolide PKS
US-09-428-517-3

```

```
Query Match          44.3%; Score 38.5; DB 4; Length 3816;
Best Local Similarity 45.0%; Pred. No. 4.4e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
```

Qy 2 ASDGALVPR- - - - - RLQHRP 16
| : | | | | : : | | |
Db 3586 AADGTLPPLLSGLVRVRHRP 3605

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,782A
FILING DATE: 28-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9710734.6
FILING DATE: 23-MAY-1997
APPLICATION NUMBER: EP 97309144.0
FILING DATE: 13-NOV-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-30167
TELECOMMUNICATION INFORMATION:
TELEPHONE: 601-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-067-782A-2

Query Match 43.7%; Score 38; DB 4; Length 351;
Best Local Similarity 57.1%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 5; Indels

Qy 3 SDGALVPRRLOHRP 16
Db 251 SSRALVPRNAQFKP 264

RESULT 7
US-09-171-461-32
; Sequence 32, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 439
; TYPE: PRT
; ORGANISM: CELO VIRUS
; FEATURE:
; OTHER INFORMATION: Position: 5366..6685/Product:IVa2
US-09-171-461-32

Query Match 43.7%; Score 38; DB 4; Length 439;
Best Local Similarity 37.5%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 8; Indels

Qy 1 CASDQALVPRRLOHRP 16

```

Db 206 CREDGTIAPKTSTFRP 221

RESULT 8

US-09-300-909-21

; Sequence 21, Application US/09300909

; Patent No. 6306580

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING

; HELICASE ACTIVITY AND METHOD THEREFOR

; NUMBER OF SEQUENCES: 27

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/300,909

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/083,942

; FILING DATE: 01-MAY-1998

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 506 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-300-909-21

Query Match

Best Local Similarity 43.7%; Score 38; DB 4; Length 506;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHRP 16

|||||:||||

Db 396 ALDGNLVSMDVKHRP 410

RESULT 9

US-08-472-666-3

; Sequence 3, Application US/08472666

; Patent No. 5821048

; GENERAL INFORMATION:

; APPLICANT: Howley, Peter M.

; APPLICANT: Benson, John D.

; APPLICANT: Yasugi, Toshiharu

; APPLICANT: Sakai, Hiroyuki

; TITLE OF INVENTION: METHOD AND KIT FOR DIAGNOSING

; PAPILLOMAVIRUS-INFECTED CELLS

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ann-Louise Kerner, Ph.D.

; ADDRESSEE: Lappin & Kusmer

; STREET: 200 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 01209

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/472,666

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: McDaniels, Patricia A.

; REGISTRATION NUMBER: 33,194

; REFERENCE/DOCKET NUMBER: HAZ-010

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-330-1300

; TELEFAX: 617-330-1311

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 649 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Human papillomavirus-16 E1

; STRAIN: HPV-16

; US-08-472-666-3

Query Match

Best Local Similarity 43.7%; Score 38; DB 2; Length 649;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHRP 16

|||||:||||

Db 539 ALDGNLVSMDVKHRP 553

RESULT 10

PCT-US96-07615-3

; Sequence 3, Application PC/TUS9607615

; GENERAL INFORMATION:

; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE

; TITLE OF INVENTION: METHODS, KITS, AND COMPOSITIONS FOR DIAGNOSING

; PAPILLOMAVIRUS INFECTION

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lappin & Kusmer

; STREET: 200 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 01209

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/07615

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Kerner, Ann-Louise

; REGISTRATION NUMBER: 33,523

; REFERENCE/DOCKET NUMBER: HAZ-010PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-330-1300

; TELEFAX: 617-330-1311

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 649 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Human papillomavirus-16 E1

; STRAIN: HPV-16

; PCT-US96-07615-3

Query Match

Best Local Similarity 43.7%; Score 38; DB 5; Length 649;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 ASDGALVPRRLQHRP 16
| | | | : | | |
Db 539 ALDGNLVSMDVKHRP 553

RESULT 11

US-08-319-866-10
; Sequence 10, Application US/08319866
; Patent No. 5929223
; GENERAL INFORMATION:
; APPLICANT: Tully, Timothy P.
; APPLICANT: Yin, Jerry C.
; APPLICANT: Regulski, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES
; TITLE OF INVENTION: ASSOCIATED WITH LONG-TERM MEMORY
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/319,866
FILING DATE: 7-OCT-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: CSHL94-03
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1205 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-319-866-10

Query Match 43.1%; Score 37.5; DB 2; Length 1205;
Best Local Similarity 69.2%; Pred. No. 1.9e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 5 GALV-PRRLQHRP 16
| | | | : | | |
Db 103 GSLVLPKRLQTRP 115

RESULT 12

US-09-123-708-6
; Sequence 6, Application US/09123708
; Patent No. 6146887
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GOEDECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2003
; CURRENT APPLICATION NUMBER: US/09/123,708
; CURRENT FILING DATE: 1998-07-28

; EARLIER APPLICATION NUMBER: 08/553,503
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: P4411402.8
; EARLIER FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1205
; TYPE: PRT
; ORGANISM: Cytomegalovirus
US-09-123-708-6

Query Match 43.1%; Score 37.5; DB 4; Length 1205;
Best Local Similarity 69.2%; Pred. No. 1.9e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 5 GALV-PRRLQHRP 16
| | | | : | | |
Db 103 GSLVLPKRLQTRP 115

RESULT 13

US-09-123-624-6
; Sequence 6, Application US/09123624
; Patent No. 6149936
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Jurgen
; APPLICANT: GOEDECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2004
; CURRENT APPLICATION NUMBER: US/09/123,624
; CURRENT FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 08/553,503
; PRIOR FILING DATE: 1996-03-01
; PRIOR APPLICATION NUMBER: 4411402.8
; PRIOR FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1205
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-123-624-6

Query Match 43.1%; Score 37.5; DB 4; Length 1205;
Best Local Similarity 69.2%; Pred. No. 1.9e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 5 GALV-PRRLQHRP 16
| | | | : | | |
Db 103 GSLVLPKRLQTRP 115

RESULT 14

US-08-858-207A-373
; Sequence 373, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA

```
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimml, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 373:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO. 6348328e
; US-08-858-207A-373

Query Match 42.5%; Score 37; DB 4; Length 137;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 LVPRRLQHR 15
Db 2 LVPRVKHR 10

RESULT 15
US-08-414-926A-26
; Sequence 26, Application US/08414926A
; Patent No. 5721354
; GENERAL INFORMATION:
; APPLICANT: Spaete, Richard
; APPLICANT: Cha, Tai-An
; TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,926A
; FILING DATE: March 31, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseir, Luan
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: AVIR-011/00US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0663
```

```
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: tol.22
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..336
; OTHER INFORMATION: /label= UL151
; US-08-414-926A-26

Query Match 42.5%; Score 37; DB 1; Length 336;
Best Local Similarity 46.7%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 CASDGALVPRRLQHR 15
Db 213 CSQHGAFFPARHLHR 227
```

Search completed: September 10, 2002, 10:45:32
Job time: 165 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 10, 2002, 10:42:47 ; Search time 12.84 Seconds
(without alignments)
30.437 Million cell updates/sec

Title: US-09-674-254-3

Perfect score: 87
Sequence: 1 CASDGLVPRRLQHRP 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	67	77.0	366	3	US-08-987-904A-2
2	67	77.0	366	3	US-08-987-904A-4
3	40	46.0	1205	1	US-07-908-245-2
4	38.5	44.3	3816	4	US-09-428-517-3
5	38	43.7	121	4	US-09-067-782A-5
6	38	43.7	351	4	US-09-067-782A-2
7	38	43.7	439	4	US-09-171-461-32
8	38	43.7	506	4	US-09-300-909-21
9	38	43.7	649	2	US-08-472-666-3
10	38	43.7	649	5	PCT-US96-07615-3
11	37.5	43.1	1205	2	US-08-319-866-10
12	37.5	43.1	1205	4	US-09-123-708-6
13	37.5	43.1	1205	4	US-09-123-624-6
14	37	42.5	137	4	US-08-858-207A-373
15	37	42.5	336	1	US-08-414-926A-26
16	37	42.5	336	2	US-08-926-922-26
17	37	42.5	336	3	US-09-253-682-26
18	37	42.5	336	4	US-09-527-657-26
19	37	42.5	395	3	US-08-981-825-6
20	37	42.5	395	4	US-09-480-784-6
21	36	41.4	95	4	US-09-142-078-62
22	36	41.4	95	4	US-09-357-141-62
23	36	41.4	924	2	US-08-588-983-18
24	36	41.4	924	2	US-08-588-976-18
25	36	41.4	2476	2	US-08-276-967-2
26	35	40.2	20	2	US-08-934-915-120
27	35	40.2	255	4	US-09-355-166-3

28 35 40.2 299 4 US-09-286-529-17 Sequence 17, Appl
29 35 40.2 300 2 US-08-794-796-2 Sequence 2, Appl
30 35 40.2 402 1 US-08-460-806-4 Sequence 4, Appl
31 35 40.2 402 1 US-08-460-806-5 Sequence 5, Appl
32 35 40.2 402 1 US-08-325-630-4 Sequence 4, Appl
33 35 40.2 402 1 US-08-325-630-5 Sequence 5, Appl
34 35 40.2 595 2 US-08-677-049-11 Sequence 11, Appl
35 35 40.2 978 2 US-08-415-593-43 Sequence 43, Appl
36 34.5 39.7 1384 4 US-08-976-255-11 Sequence 5, Appl
37 34 39.1 16 1 US-08-307-724B-5 Sequence 5, Appl
38 34 39.1 159 3 US-08-390-353A-5 Sequence 7, Appl
39 34 39.1 173 4 US-09-095-758-7 Sequence 7, Appl
40 34 39.1 173 4 US-09-422-968-7 Sequence 4, Appl
41 34 39.1 175 4 US-09-028-328-4 Sequence 4, Appl
42 34 39.1 210 1 US-08-320-559-4 Sequence 4, Appl
43 34 39.1 210 1 US-08-327-392-4 Sequence 4, Appl
44 34 39.1 210 3 US-08-545-860D-4 Sequence 4, Appl
45 34 39.1 210 5 PCT-US94-04496-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-987-904A-2

; Sequence 2, Application US/08987904A

; Patent No. 6027917

; GENERAL INFORMATION:

; APPLICANT: Celeste, Anthony J.

; APPLICANT: Murray, Beth

; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN (BMP) - 17 AND BMP-18

; TITLE OF INVENTION: COMPOSITIONS

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genetics Institute, Inc.

; STREET: 87 Cambridgepark Drive

; CITY: Cambridge

; STATE: MA

; COUNTRY: US

; ZIP: 02140

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/987,904A

; FILING DATE: 10-DEC-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: LAZAR, STEVEN R

; REGISTRATION NUMBER: 32,618

; REFERENCE/DOCKET NUMBER: GI 5307

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 498-8769

; TELEFAX: (617) 876-8581

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 366 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-987-904A-2

Query Match 77.0%; Score 67; DB 3; Length 366;

Best Local Similarity 100.0%; Pred. No. 0.00024;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CASDGLVPRRLQ 13

|||||

Db 353 CASDGLVPRRLQ 365

```

RESULT 2
US-08-987-904A-4
; Sequence 4, Application US/08987904A
; Patent No. 6027917
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Murray, Beth
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN (BMP)- 17 AND BMP-18
; TITLE OF INVENTION: COMPOSITIONS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: US
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/987,904A
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, STEVEN R
; REGISTRATION/DOCKET NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: GI 5307
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8769
; TELEFAX: (617) 876-8581
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-987-904A-4

Query Match 77.0%; Score 67; DB 3; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CASDQALVPRRLQ 13
Db 353 CASDQALVPRRLQ 365

3ULT 3
-07-908-245-2
; Sequence 2, Application US/07908245
; Patent No. 5498539
; GENERAL INFORMATION:
; APPLICANT: Harrison, David G.
; APPLICANT: Alexander, R. Wayne
; APPLICANT: Murphy, T.J.
; APPLICANT: Nishida, Ken'ichi
; TITLE OF INVENTION: Endothelial Nitric Oxide Synthase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,245
; FILING DATE: 19920702
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION/DOCKET NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; TISSUE TYPE: Aorta
; CELL TYPE: Endothelial
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 496..512
; OTHER INFORMATION: /note= "CA++/CAM binding domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 651..678
; OTHER INFORMATION: /note= "FMN binding domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 795..806
; OTHER INFORMATION: /note= "FAD-Pyrophosphate binding
; OTHER INFORMATION: domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 937..947
; OTHER INFORMATION: /note= "FAD-Isolalloxanthine
; OTHER INFORMATION: binding domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1012..1030
; OTHER INFORMATION: /note= "NADPH-Ribose binding
; OTHER INFORMATION: domain"
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1111..1124
; OTHER INFORMATION: /note= "NADPH-Ribose binding
; OTHER INFORMATION: domain"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 33..34
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 46..47
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 53..54
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 58..59

```


Query Match 77.0%; Score 67; DB 20; Length 120;
 Best Local Similarity 100.0%; Pred. No. 0.00024;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQ 13
 DB 107 casdglvprrlq 119
 |||||

RESULT 6
 AAY17870
 ID AAY17870 standard; Protein; 366 AA.
 AC AAY17870;

DT 20-AUG-1999 (first entry)

DE Human bone morphogenic protein BMP-17.

KW Human; bone morphogenic protein; BMP-17; BMP-18; cartilage; tendon;
 KW connective tissue defect; ligament; meniscus; wound healing; growth;
 KW differentiation; epidermis; muscle; nerve; cardiac muscle.

OS Homo sapiens.

TH Key Location/Qualifiers
 FT Peptide 1..142
 FT Protein /label= pro-peptide
 FT 143..366
 FT /label= BMP-17

PN W09929718-A2.

PD 17-JUN-1999.

PF 17-NOV-1998; 98WO-US24613.

PR 10-DEC-1997; 97US-0987904.

XX (GEMY) GENETICS INST INC.

PI Celeste AJ, Murray BL;

DR WPI; 1999-385570/32.

DR N-PSDB; AAX80213.

XX New Purified bone morphogenic protein-17 and -18 (BMP-17 and BMP-18)
 PT polypeptides, useful for the induction of growth and/or
 PT differentiation of undifferentiated embryonic and stem cells

PS Claim 16; Page 35-36; 39pp; English.

CC The present sequence is a human bone morphogenic protein designated
 CC BMP-17. BMP proteins are useful for the induction of growth and/or
 CC differentiation of undifferentiated embryonic and stem cells, and for
 CC the treatment of bone, cartilage and other connective tissue defects
 CC including tendons, ligaments and meniscus, in wound healing and related
 CC tissue repair, and for treatment of disorders and defects to tissues
 CC which include epidermis, nerve, muscle, including cardiac muscle, and
 CC other tissues and wounds, and organs such as liver, lung, epithelium,
 CC brain, spleen, cardiac, pancreas and kidney tissue. DNA encoding BMP
 CC proteins can be useful as probes to detect expression of BMP proteins,
 CC and the vectors containing DNA encoding BMP proteins are useful for
 CC delivery of the BMP proteins to cells of a patient.

XX Sequence 366 AA;

Query Match 77.0%; Score 67; DB 20; Length 366;
 Best Local Similarity 100.0%; Pred. No. 0.00078;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQ 13

DB 353 casdglvprrlq 365
 |||||

RESULT 7
 AAY17871

ID AAY17871 standard; Protein; 366 AA.

AC AAY17871;

DT 20-AUG-1999 (first entry)

DE Human bone morphogenic protein BMP-18.

KW Human; bone morphogenic protein; BMP-17; BMP-18; cartilage; tendon;
 KW connective tissue defect; ligament; meniscus; wound healing; growth;
 KW differentiation; epidermis; muscle; nerve; cardiac muscle.

OS Homo sapiens.

TH Key Location/Qualifiers
 FT Peptide 1..135
 FT Protein /label= pro-peptide
 FT 136..366
 FT /label= BMP-18

PN W09929718-A2.

PD 17-JUN-1999.

PF 17-NOV-1998; 98WO-US24613.

PR 10-DEC-1997; 97US-0987904.

XX (GEMY) GENETICS INST INC.

PI Celeste AJ, Murray BL;

DR WPI; 1999-385570/32.

DR N-PSDB; AAX80214.

XX New Purified bone morphogenic protein-17 and -18 (BMP-17 and BMP-18)
 PT polypeptides, useful for the induction of growth and/or
 PT differentiation of undifferentiated embryonic and stem cells

PS Claim 20; Page 38-39; 39pp; English.

CC The present sequence is a human bone morphogenic protein designated
 CC BMP-18. BMP proteins are useful for the induction of growth and/or
 CC differentiation of undifferentiated embryonic and stem cells, and for
 CC the treatment of bone, cartilage and other connective tissue defects
 CC including tendons, ligaments and meniscus, in wound healing and related
 CC tissue repair, and for treatment of disorders and defects to tissues
 CC which include epidermis, nerve, muscle, including cardiac muscle, and
 CC other tissues and wounds, and organs such as liver, lung, epithelium,
 CC brain, spleen, cardiac, pancreas and kidney tissue. DNA encoding BMP
 CC proteins can be useful as probes to detect expression of BMP proteins,
 CC and the vectors containing DNA encoding BMP proteins are useful for
 CC delivery of the BMP proteins to cells of a patient.

XX Sequence 366 AA;

Query Match 77.0%; Score 67; DB 20; Length 366;
 Best Local Similarity 100.0%; Pred. No. 0.00078;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQ 13

DB 353 casdglvprrlq 365
 |||||

RESULT 8

```

AAV13363
ID AAY13363 standard; Protein; 366 AA.
XX
AC AAY13363;
XX
DT 25-JUN-1999 (first entry)
XX
DE Amino acid sequence of protein PRO317.
XX
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophla areata;
KW anti-thrombotic; wound healing; tissue repair.
XX
OS Homo sapiens.
XX
PN WO9914328-A2.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US19330.
XX
PR 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063341.
PR 28-OCT-1997; 97US-0063342.
PR 28-OCT-1997; 97US-0063344.
PR 28-OCT-1997; 97US-0063349.
PR 28-OCT-1997; 97US-0063350.
PR 28-OCT-1997; 97US-0063356.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.
PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.

PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.
XX
PA (GETH ) GENENTECH INC.
XX
PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
XX
DR WPI; 1999-229533/19.
DR N-PSDB; AAX52234.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
XX gastrointestinal ulceration
PS Claim 12; Fig 42; 320pp; English.
XX
CC AAY13344-403 represent secreted and transmembrane human proteins.
CC The cDNA sequences are obtained from cDNA libraries, prepared from
CC fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
CC The encoded polypeptides have specific uses based on their homology to
CC known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
CC associated with the preservation and maintenance of gastrointestinal
CC mucosa and the repair of acute and chronic mucosal lesions
CC (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
CC ulceration and congenital microvillus atrophy), skin diseases associated
CC with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
CC cancers such as lung squamous cell carcinoma of the vulva and gliomas),
CC potent effects on cell growth and development, diseases related to
CC growth or survival of nerve cells including Parkinson's disease,
CC Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
CC for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
CC of Usher Syndrome or Atrophla areata; PRO269 can be used as an
CC anti-thrombotic agent; PRO287 polypeptides and portions may have
CC therapeutic applications in wound healing and tissue repair; PRO317 can
CC be used for treating problems of the kidney, uterus, endometrium, blood
CC vessels, or related tissue, e.g. in the heart of genital tract.
XX
SQ Sequence 366 AA;

Query Match 77.0%; Score 67; DB 20; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CASDGLVPRRLQ 13
|||||
Db 353 casdglvprrlq 365

RESULT 9
AAY03850
ID AAY03850 standard; Protein; 366 AA.
XX
AC AAY03850;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human lefty protein.
XX
KW Nodal protein; lefty protein; TGF-beta; sexual development; human;
KW pituitary; cartilage; osteoarthritis; osteoporosis; haematopoiesis;
KW periodontal disease; wound healing; tissue repair; tumour; cancer;
KW interstitial lung disease; autoimmunity; leukaemia; lymphoma; immunity;
KW immunosuppression; inflammatory bowel disease; myelosuppression;
KW infectious disease; bone.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Peptide 1..18 /note= "signal peptide"
FT Protein 19..366 /note= "mature protein"

```

```

FT FT      78..364
FT /note= "first predicted TGF-beta like domain of lefty"
FT Domain 136...366
FT /note= "second predicted TGF-beta like domain of lefty"
FT Domain 143...366
FT /note= "third predicted TGF-beta like domain of lefty"
XX XX
PN W09509198-A1.
XX
XX 25-FEB-1999.
XX
XX 20-AUG-1998; 98WO-US17211.
XX
XX 21-AUG-1997; 97US-0056565.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Ruben SM, Soppet DR;
XX
XX WPI: 1999-190173/16.
XX N-PSDB; AAX31925.
XX
XX New isolate human Nodal and Lefty polypeptides
XX
XX Claim 1: Fig 1B; 182pp; English.
XX
XX The present invention relates to novel human nodal and lefty proteins
XX which are members of the TGF-beta family. The human nodal and lefty
XX proteins may be involved in a developmental process such as the correct
XX formation of various structures or in one or more post-developmental
XX capacities including sexual development, pituitary hormone production,
XX and the creation of bone and cartilage. The Nodal and Lefty polypeptides
XX are useful for enhancing or enriching the growth and/or differentiation
XX of specific cell populations, eg. embryonic cells or stem cells. They can
XX be used to treat such conditions as osteoarthritis, osteoporosis, and
XX other abnormalities of bone, cartilage, muscle, tendon, ligament, and/or
XX other connective tissues and/or organs such as liver, lung, cardiac,
XX pancreas, and kidney. Compositions containing nodal and lefty proteins
XX may be useful for growth formation, for treating periodontal disease and
XX for modulating haematopoiesis, wound healing and tissue repair. They can
XX also be used for the treatment of tumours, cancers, interstitial lung
XX disease, and any dysregulation of the growth and differentiation patterns
XX of cell function including autoimmunity, arthritis, leukaemia, lymphomas,
XX immunosuppression, immunity, humoral immunity, inflammatory bowel
XX disease, myelosuppression, or infectious diseases. The present sequence
XX represents a human lefty polypeptide. The cDNA encoding the lefty
XX protein is deposited under the ATCC deposit No. 209091.
XX
XX Sequence 366 AA;
SQ
Query Match 77.0%; Score 67; DB 20; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
<Y 1 CASDGLVPRRLQ 13
Db 353 casdglvprrlq 365
|||||
RESULT 10
ID AAY05287 standard; Protein; 366 AA.
XX
XX AAY05287;
XX
XX 22-JUN-1999 (first entry)
XX
XX EGF-like homologue EBAF-2.
XX
XX Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;
XX EBAF-2; inhibitor; tumour growth; cancer; EGF-like homologue;
XX FGF-8 homologue.
XX

```

```

XX Homo sapiens.
XX OS
XX WO9914327-A2.
XX PN
XX 25-MAR-1999.
XX PD
XX 10-SEP-1998; 98WO-US18824.
XX PF
XX 25-NOV-1997; 97US-0066840.
XX PR
XX 17-SEP-1997; 97US-0059114.
XX PR
XX 17-SEP-1997; 97US-0059117.
XX PR
XX 18-SEP-1997; 97US-0059263.
XX PR
XX 15-OCT-1997; 97US-0062125.
XX PR
XX 17-OCT-1997; 97US-0062285.
XX PR
XX 17-OCT-1997; 97US-0062287.
XX PR
XX 24-OCT-1997; 97US-0062816.
XX PR
XX 29-OCT-1997; 97US-0063704.
XX
XX (GETH ) GENENTECH INC.
XX PA
XX Botstein D, Goddard A, Gurney A, Hillan K, Lawrence DA;
XX PI
XX Roy M, Wood WI;
XX
XX WPI: 1999-229532/19.
XX DR
XX N-PSDB; AAX28437.
XX
XX Antibodies against specific proteins overexpressed in tumours
XX
XX Example 1: Fig 30; 130pp; English.
XX
XX This sequence represents the EGF-like homologue EBAF-2.
XX The invention relates to antibodies (Ab) that bind to any of the
XX polypeptides (I) designated PRO187; PRO533; PRO214; PRO240; PRO211;
XX PRO230; PRO261; PRO246 or EBAF-2. The Ab, or other agents that inhibit
XX expression and/or activity of (I) are used: (i) to inhibit growth of
XX tumours; and (ii) as diagnostic/prognostic reagents for detection or
XX quantification of (I) in cells or tissues, by standard immunoassays, with
XX overexpression being indicative of cancer. For therapeutic use, the Ab
XX may be conjugated to a toxin, chemotherapeutic agent or radioisotope.
XX Genes expressing (I), many of which are growth factor homologues, are
XX overexpressed in some cases of cancer.
XX
XX Sequence 366 AA;
SQ
Query Match 77.0%; Score 67; DB 20; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CASDGLVPRRLQ 13
Db 353 casdglvprrlq 365
|||||
RESULT 11
ID AAY88575 standard; Protein; 366 AA.
XX
XX AAY88575;
XX
XX 09-AUG-2000 (first entry)
XX
XX Human PRO317 amino acid sequence.
XX
XX Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;
XX PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;
XX cell growth; proliferation; transforming growth factor; ADEPT;
XX antibody dependent enzyme mediated prodrug therapy.
XX
XX Homo sapiens.
XX OS
XX WO200015666-A2.
XX PN

```



```

PR 26-JUL-1999; 99US-0145698.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 02-DEC-1999; 99WO-US28564.
XX (GETH ) GENENTECH INC.
PA
XX
XX Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;
PI WPI; 2001-091968/10.
DR N-PSDB; AAF60376.
XX
XX New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,
PT useful for diagnosing and treating cancers -
PT
XX Claim 61; Fig 18; 196pp; English.
XX
XX The present invention relates to PRO proteins and coding sequences. The
CC present sequence is one such PRO protein. It was found that the PRO genes
CC are amplified in the genome of tumour cells. The gene amplification is
CC expected to be associated with the overexpression of the gene product and
CC contributes to tumourigenesis. Therefore, antagonists of PRO proteins are
CC useful for the treatment of benign or malignant tumours, leukaemias,
CC lymphoid malignancies and other disorders such as neuronal, glial,
CC astrocytal, hypothalamic, glandular, epithelial, inflammatory and
CC immunologic disorders.
XX
XX Sequence 366 AA;
SQ
Query Match 77.0%; Score 67; DB 22; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CASDGLVPRRLQ 13
DB 353 casdgalvprrlq 365
RESULT 14
AAB80231
ID AAB80231 standard; Protein; 366 AA.
AC AAB80231;
XX
XX 24-APR-2001 (first entry)
XX Human PRO317 protein.
XX
XX Human; PRO; dermatologic; antipsoriatic; cytostatic; antiinflammatory;
XX antiparkinsonian nootropic; neuroprotective; vulneryary; cardiant;
XX antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
XX antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;
XX ophthalmologic; gene therapy; skin disease; gastrointestinal disorder;
XX ischaemia; inflammation.
XX
XX Homo sapiens.
XX
XX WO200104311-A1.
XX
XX 18-JAN-2001.
XX
XX 22-FEB-2000; 2000WO-US04414.
XX
XX 07-JUL-1999; 99US-0143048.
XX 26-JUL-1999; 99US-0145698.
XX 28-JUL-1999; 99US-0146222.
XX 08-SEP-1999; 99WO-US20594.
XX 13-SEP-1999; 99WO-US20944.

```

```

PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30995.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 99WO-US00219.
XX (GETH ) GENENTECH INC.
PA
XX
XX Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX WPI; 2001-081051/09.
XX N-PSDB; AAF72392.
XX
XX Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease) -
XX
XX Claim 1; Fig 42; 393pp; English.
XX
XX The present sequence is one of sixty one novel secreted and
CC transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemias such as coronary
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC diabetes and retinal disorders such as retinitis pigmentosum.
CC The PRO nucleic acids have applications in molecular biology, including
CC use as hybridization probes, and in chromosome and gene mapping.
XX
XX Sequence 366 AA;
SQ
Query Match 77.0%; Score 67; DB 22; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CASDGLVPRRLQ 13
DB 353 casdgalvprrlq 365
RESULT 15
AAB19837
ID AAB19837 standard; Protein; 366 AA.
XX
XX AAB19837;
XX
XX 05-MAR-2001 (first entry)
XX
XX Endometrial bleeding associated factor (ebaf) protein.
XX
XX Endometrial bleeding associated factor; ebaf; human;
XX chromosome 1q42.1; neurodegenerative disease; Alzheimer's disease;
XX Parkinson's disease; Huntington's chorea; dementia;
XX amyotrophic lateral sclerosis; Pick's disease; therapy.
XX
XX Homo sapiens.
XX
XX WO20006068-A2.
XX
XX 09-NOV-2000.

```

```

XX 28-APR-2000; 2000WO-US11623.
XX
XX 29-APR-1999; 99US-0302094.
XX
XX (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.
XX
XX Tabibzadeh S;
XX
XX WPI; 2001-040876/05.
XX
XX N-PSDB; AAA88913.
XX
XX Inducing growth and enhancing survival of nervous tissue by contacting
XX with endometrial bleeding associated factor protein
XX
XX Disclosure; Fig 2; 23pp; English.
XX
XX The present sequence is that of human endometrial bleeding
XX associated factor (ebaf). A claimed method for inducing growth
XX and enhancing survival of nervous tissue comprises contacting the
XX tissue with ebaf or a nucleic acid encoding ebaf in vitro or in
XX vivo. The method can be used to treat damaged or degenerated
XX nervous tissue resulting from injury associated with trauma,
XX diabetes, kidney dysfunction, ischaemia or use of therapeutic
XX agents, or to treat a neurodegenerative disease such as Alzheimer's
XX disease, Parkinson's disease, Huntington's chorea, amyotrophic
XX lateral sclerosis, dementia, or Pick's disease (all claimed).
XX
SQ Sequence 366 AA;

Query Match          77.0%; Score 67; DB 22; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CASDGLVPRRLQ 13
   | | | | | | | | | |
Db 353 casdgalvprrlq 365

```

Search completed: September 10, 2002, 10:45:12
Job time: 265 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	44	50.6	67	4	Q9UHN8	Q9uhn8	homo sapien
2	44	50.6	99	4	Q9UI41	Q9ui41	homo sapien
3	44	50.6	120	4	Q9UHW7	Q9uhw7	homo sapien
4	44	50.6	152	4	Q9UR55	Q9ubr5	homo sapien
5	44	50.6	219	12	Q7I131	Q7i131	prunus necr
6	44	50.6	224	12	Q9YKE6	Q9yke6	prunus necr
7	44	50.6	224	12	Q9YKE3	Q9yke3	prunus necr
8	44	50.6	224	12	Q70915	Q70915	prunus necr
9	44	50.6	224	12	Q70917	Q70917	prunus necr
10	44	50.6	224	12	Q9IMR7	Q9imr7	prunus necr
11	44	50.6	224	12	Q9IMR5	Q9imr5	prunus necr
12	44	50.6	224	12	Q9IMR2	Q9imr2	prunus necr
13	44	50.6	224	12	Q9IMQ9	Q9imq9	prunus necr
14	44	50.6	224	12	Q9IMQ7	Q9imq7	prunus necr
15	44	50.6	224	12	Q9IMQ3	Q9imq3	prunus necr
16	44	50.6	224	12	Q9IMQ1	Q9imq1	prunus necr


```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21308461; PubMed=11415443;
RA Han W., Lou Y., Tang J., Zhang Y., Chen Y., Li Y., Gu W., Huang J.,
RA Gu L., Tang Y., Li F., Song Q., Di C., Wang L., Shi Q., Sun R.,
RA Xia D., Rui M., Tang J., Ma D.;
RT "Molecular cloning and characterization of chemokine-like factor 1
RT (CKLF1), a novel human cytokine with unique structure and potential
RT chemotactic activity.";
RL Biochem. J. 357:127-135(2001).
DR EMBL; AF096895; AAF06722.1; -.
SQ SEQUENCE 99 AA; 10923 MW; 1F5EBF5AA7A9479E CRC64;

Query Match 50.6%; Score 44; DB 4; Length 99;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHRP 16
DB 68 CLADGALIVKLLFNP 83
I :||||: |:-|

RESULT 3
Q9UHM7 PRELIMINARY; PRT; 120 AA.
AC Q9UHM7;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE CHEMOKINE-LIKE FACTOR 4.
GN CKLF4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Han W.L.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF145216; AAF19350.1; -.
SQ SEQUENCE 120 AA; 13899 MW; BF0C011077F24C13 CRC64;

Query Match 50.6%; Score 44; DB 4; Length 120;
Best Local Similarity 50.0%; Pred. No. 3.7;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHRP 16
DB 89 CLADGALIVKLLFNP 104
I :||||: |:-|

RESULT 4
Q9UBR5 PRELIMINARY; PRT; 152 AA.
AC Q9UBR5;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE TRANSMEMBRANE PROTEOLIPID (HSPC224) (UNKNOWN) (PROTEIN FOR
DE MGC:10558).
GN C32 OR CKLF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=PANCREAS;
RA Zhang J.S., Nelson M., Wang L., Smith D.I.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Han W., Gu W., Li Y., Zhang Y., Song Q., Di C., Ma D.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL CORD BLOOD;
RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
RT "Human full length cDNA cloned from cd34+ stem cells.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREATIC ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF057306; AAF21255.1; -.
DR EMBL; AF135380; AAF19599.1; -.
DR EMBL; AF151058; AAF36144.1; -.
DR EMBL; BC004380; AAH04380.1; -.
KW Transmembrane.
SQ SEQUENCE 152 AA; 17170 MW; 6149D6D2B6AED9F3 CRC64;

Query Match 50.6%; Score 44; DB 4; Length 152;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHRP 16
DB 121 CLADGALIVKLLFNP 136
I :||||: |:-|

RESULT 5
O71131 PRELIMINARY; PRT; 219 AA.
ID O71131;
AC O71131;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-AUG-1998 (TremBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
OS Prunus necrotic ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Ilarvirus.
OX NCBI_TaxID=37733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH61;
RX MEDLINE=98343741; PubMed=9680147;
RA Hammond R.W., Crosslin J.M.;
RT "Virulence and molecular polymorphism of Prunus necrotic ringspot
RT virus isolates.";
RL J. Gen. Virol. 79:1815-1823(1998).
DR EMBL; AF034989; AAC41025.1; -.
DR InterPro; IPR002681; Ilar_coat.
DR Pfam; PF01787; Ilar_coat; 1.
FT NON_TER 219
FT SEQUENCE 219 AA; 24491 MW; 0B2748B65C5DE4D5 CRC64;

Query Match 50.6%; Score 44; DB 12; Length 219;
Best Local Similarity 60.0%; Pred. No. 6.8;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHR 15
DB 20 CHPN GALVPLRAQOR 34
I :||||: |:-|

RESULT 6
```

```

Q9YKE6
ID Q9YKE6 PRELIMINARY; PRT; 224 AA.
AC Q9YKE6;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE COAT PROTEIN (CAPSID PROTEIN).
GN CP.
OS Prunus necrotic ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Ilarvirus.
OX NCBI_TaxID=37733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCIT.MAY1;
RA Aparicio F., Myrta A., Di Terlizzi B., Pallas V.;
RT "Molecular variability among fifteen isolates of Prunus necrotic ringspot virus (PNRSV) from six different Prunus species.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=4/8, AND UN;
RX MEDLINE=20349238; PubMed=10893149;
RA Vaskova D., Petrzik K., Karesova R.;
RT "Variability and molecular typing of the woody-tree infecting prunus necrotic ringspot ilarvirus.";
RL Arch. Virol. 145:699-709(2000).
DR EMBL; AJ133205; CAB37307.1; -
DR EMBL; AF170165; AAF89719.1; -
DR EMBL; AF170163; AAF89715.1; -
DR InterPro; IPR002681; Ilar_coat.
DR Pfam; PF01787; Ilar_coat; 1.
KW Coat protein.
SQ SEQUENCE 224 AA; 24911 MW; DE00E9AED5E4E75D CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 CASDGVLPRLQHR 15
DB 20 CHPNGALVPLRAQOR 34
I :||||| I I

RESULT 7
ID Q9YKE3 PRELIMINARY; PRT; 224 AA.
AC Q9YKE3;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE COAT PROTEIN (CAPSID PROTEIN).
GN CP.
OS Prunus necrotic ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Ilarvirus.
OX NCBI_TaxID=37733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTSP.MURI;
RA Aparicio F., Myrta A., Di Terlizzi B., Pallas V.;
RT "Molecular variability among fifteen isolates of Prunus necrotic ringspot virus (PNRSV) from six different Prunus species.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6/54;
RX MEDLINE=20349238; PubMed=10893149;
RA Vaskova D., Petrzik K., Karesova R.;
RT "Variability and molecular typing of the woody-tree infecting prunus necrotic ringspot ilarvirus.";
RL Arch. Virol. 145:699-709(2000).

```

```

DR EMBL; AJ133208; CAB37310.1; -
DR EMBL; AF170160; AAF89709.1; -
DR InterPro; IPR002681; Ilar_coat.
DR Pfam; PF01787; Ilar_coat; 1.
KW Coat protein.
SQ SEQUENCE 224 AA; 24927 MW; E215BCEED1C0703F CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 CASDGVLPRLQHR 15
DB 20 CHPNGALVPLRAQOR 34
I :||||| I I

RESULT 8
ID Q70915 PRELIMINARY; PRT; 224 AA.
AC Q70915;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE COAT PROTEIN.
GN 3B.
OS Prunus necrotic ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Ilarvirus.
OX NCBI_TaxID=37733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MISSION;
RA Scott S.W., Zimmerman M.T., Ge X., Mackenzie D.J.;
RT "The coat proteins and putative movement proteins of isolates of Prunus necrotic ringspot virus from different host species and geographic origins are extensively conserved.";
RL Eur. J. Plant Pathol. 104:155-161(1998).
DR EMBL; AF013285; AAC16499.1; -
DR InterPro; IPR002681; Ilar_coat.
DR Pfam; PF01787; Ilar_coat; 1.
SQ SEQUENCE 224 AA; 25088 MW; CA6FB9979E4FCA87 CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 CASDGVLPRLQHR 15
DB 20 CHPNGALVPLRAQOR 34
I :||||| I I

RESULT 9
ID Q70917 PRELIMINARY; PRT; 224 AA.
AC Q70917;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE COAT PROTEIN.
GN 3B.
OS Prunus necrotic ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Ilarvirus.
OX NCBI_TaxID=37733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PRUNE;
RA Scott S.W., Zimmerman M.T., Ge X., Mackenzie D.J.;
RT "The coat proteins and putative movement proteins of isolates of Prunus necrotic ringspot virus from different host species and geographic origins are extensively conserved.";

```

RL Eur. J. Plant Pathol. 104:155-161(1998).
 DR EMBL: AF013286; AAC16501.1; -.
 DR InterPro: IPR002681; Iilar_coat.
 DR Pfam: PF01787; Iilar_coat; 1.
 SQ SEQUENCE 224 AA; 24939 MW; 8D03DAAEDC21BACF CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
 Best Local Similarity 60.0%; Pred. No. 7;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHR 15
 | :||||| | | |
 Db 20 CYPNGALVPLRAQQR 34

RESULT 10

ID Q9IMR7 PRELIMINARY; PRT; 224 AA.
 AC Q9IMR7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CAPSID PROTEIN.
 GN CP.
 OS Prunus necrotic ringspot virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
 OC Iilarvirus.
 OX NCBI_TaxID=37733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PS 7/11;
 RX MEDLINE=20349238; PubMed=10893149;
 RA Vaskova D., Petrzik K., Karesova R.;
 RT "Variability and molecular typing of the woody-tree infecting prunus
 RT necrotic ringspot Iilarvirus.";
 RL Arch. Virol. 145:699-709(2000).
 DR EMBL: AF170161; AAF89711.1; -.
 DR InterPro: IPR002681; Iilar_coat.
 DR Pfam: PF01787; Iilar_coat; 1.
 SQ SEQUENCE 224 AA; 24965 MW; 9D1EB6CBC04DDFDA CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
 Best Local Similarity 60.0%; Pred. No. 7;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHR 15
 | :||||| | | |
 Db 20 CYPNGALVPLRAQQR 34

RESULT 11

ID Q9IMR5 PRELIMINARY; PRT; 224 AA.
 AC Q9IMR5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CAPSID PROTEIN.
 GN CP.
 OS Prunus necrotic ringspot virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
 OC Iilarvirus.
 OX NCBI_TaxID=37733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PS 7/12;
 RX MEDLINE=20349238; PubMed=10893149;
 RA Vaskova D., Petrzik K., Karesova R.;
 RT "Variability and molecular typing of the woody-tree infecting prunus
 RT necrotic ringspot Iilarvirus.";
 RL Arch. Virol. 145:699-709(2000).

DR EMBL: AF170162; AAF89713.1; -.
 DR InterPro: IPR002681; Iilar_coat.
 DR Pfam: PF01787; Iilar_coat; 1.
 SQ SEQUENCE 224 AA; 24906 MW; F12975BF70C89ED CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
 Best Local Similarity 60.0%; Pred. No. 7;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHR 15
 | :||||| | | |
 Db 20 CYPNGALVPLRAQQR 34

RESULT 12

ID Q9IMR2 PRELIMINARY; PRT; 224 AA.
 AC Q9IMR2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CAPSID PROTEIN.
 GN CP.
 OS Prunus necrotic ringspot virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
 OC Iilarvirus.
 OX NCBI_TaxID=37733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-7/20;
 RX MEDLINE=20349238; PubMed=10893149;
 RA Vaskova D., Petrzik K., Karesova R.;
 RT "Variability and molecular typing of the woody-tree infecting prunus
 RT necrotic ringspot Iilarvirus.";
 RL Arch. Virol. 145:699-709(2000).
 DR EMBL: AF170164; AAF89717.1; -.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR002681; Iilar_coat.
 DR Pfam: PF01787; Iilar_coat; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 SQ SEQUENCE 224 AA; 24882 MW; 5C2CCB6CF2AFBA71 CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
 Best Local Similarity 60.0%; Pred. No. 7;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASD GALVPRRLQHR 15
 | :||||| | | |
 Db 20 CYPNGALVPLRAQQR 34

RESULT 13

ID Q9IMQ9 PRELIMINARY; PRT; 224 AA.
 AC Q9IMQ9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CAPSID PROTEIN.
 GN CP.
 OS Prunus necrotic ringspot virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
 OC Iilarvirus.
 OX NCBI_TaxID=37733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PS 7/5A;
 RX MEDLINE=20349238; PubMed=10893149;
 RA Vaskova D., Petrzik K., Karesova R.;
 RT "Variability and molecular typing of the woody-tree infecting prunus
 RT necrotic ringspot Iilarvirus.";

RL Arch. Virol. 145:699-709(2000).
DR EMBL; AF170166; AAF89721.1; -.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR002681; Ilar_coat.
DR Pfam; PF01787; Ilar_coat; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
SQ SEQUENCE 224 AA; 24880 MW; 9401E4CEC3222B21 CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQHR 15
| :||||| | | |
Db 20 CHPNGALVPLRAQR 34

RESULT 14
Q9IMQ7 PRELIMINARY; PRT; 224 AA.
AC Q9IMQ7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CAPSID PROTEIN.
GN CP.
OS Prunus necrotic ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Iilarvirus.
OX NCBI_TaxID=37733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UHI;
RX MEDLINE=20349238; PubMed=10893149;
RA Vaskova D., Petrzik K., Karesova R.;
RT "Variability and molecular typing of the woody-tree infecting prunus
necrotic ringspot ilarvirus";
RL Arch. Virol. 145:699-709(2000).
DR EMBL; AF170167; AAF89723.1; -.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR002681; Ilar_coat.
DR Pfam; PF01787; Ilar_coat; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
SQ SEQUENCE 224 AA; 24879 MW; 28F24B0801EC4D7D CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQHR 15
| :||||| | | |
Db 20 CHPNGALVPLRAQR 34

RESULT 15
Q9IMQ3 PRELIMINARY; PRT; 224 AA.
AC Q9IMQ3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CAPSID PROTEIN.
GN CP.
OS Prunus necrotic ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Iilarvirus.
OX NCBI_TaxID=37733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NA HRBU;
RX MEDLINE=20349238; PubMed=10893149;

RA Vaskova D., Petrzik K., Karesova R.;
RT "Variability and molecular typing of the woody-tree infecting prunus
necrotic ringspot ilarvirus";
RL Arch. Virol. 145:699-709(2000).
DR EMBL; AF170170; AAF89729.1; -.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR002681; Ilar_coat.
DR Pfam; PF01787; Ilar_coat; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
SQ SEQUENCE 224 AA; 24863 MW; 2DE64B0811ED40A7 CRC64;

Query Match 50.6%; Score 44; DB 12; Length 224;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CASDGLVPRRLQHR 15
| :||||| | | |
Db 20 CHPNGALVPLRAQR 34

Search completed: September 10, 2002, 10:46:25
Job time: 123 sec